TRANSITIONAL EMBEDDED INSTRUCTIONS FOR MANIPULATING PHYSICAL OBJECTS

A Dissertation
Presented to
The Academic Faculty

by

Keith R Bujak

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Psychology

Georgia Institute of Technology
May 2014

TRANSITIONAL EMBEDDED INSTRUCTIONS FOR MANIPULATING PHYSICAL OBJECTS

Approved by:

Dr. Richard Catrambone, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Frank Durso
School of Psychology
Georgia Institute of Technology

Dr. Bruce Walker
School of Psychology
Georgia Institute of Technology

Dr. Blair MacIntyre
School of Interactive Computing
Georgia Institute of Technology

Dr. Michael Schatz
School of Physics
Georgia Institute of Technology

Date Approved: April 2, 2014

ACKNOWLEDGEMENTS

I wish to thank Richard Catrambone and my committee for teaching me the value of asking good questions.

I wish to thank my lab mates, both past and present. Elsa Eiríksdóttir always expected me to be precise and clear with my language. Lauren Margulieux was my constant statistics sounding board.

I wish to thank Greg Gunn, who turned my system specifications into a reality.

I wish to thank the many undergraduate research assistants, without whom it would have been nearly impossible to collect and code all of my data.

I wish to thank Kathi Olson, Jenay Beer, and Sara McBride, all of whom made graduate school a richly fulfilling experience.

And finally, I wish to thank my parents, who were (and are) always there for me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

There has been much research on how people use instructional information to gain procedural knowledge. In the context of procedures involving physical objects, however, there has been little research on the role these objects play in conveying procedural information. This study investigated how people used instructions – presented as either images or text – to assemble various physical objects. Objects were selected that either comprised uniquely shaped or interchangeable parts. Participants assembled each object twice, randomly receiving either image or text instructions for each build. They then assembled each object without the instructions and made judgments about the order of the procedure from memory. Image instructions generally resulted in faster and more accurate assemblies as well as more accurate memory for procedural order. These results were found only for objects with uniquely shaped parts. An object comprising interchangeable parts was readily assembled with either instructional type. Although text alone failed to provide any advantages, the combination of images and then text resulted in more consistent mental workload, which might be beneficial in some operational contexts. These results provide insights about how physical objects influence the use of and knowledge gained from procedural instructions.

CHAPTER 1

INTRODUCTION

People regularly rely on procedural knowledge to interact with the physical world and accomplish goals. Instructions have, for many years and in many research experiments, been conveyed linguistically and graphically (Höffler & Leutner, 2007). Instructional designers make decisions about the representation, such as text or images, through which to convey this procedural information. The choice of representation influences a person's ability to follow or remember the information presented in the instructions. Although there have been many studies comparing text, images, and combinations of the two (e.g., Brunyé, Taylor, & Rapp, 2008; Konz & Dickey, 1969; Larkin & Simon, 1987; Palmiter & Elkerton, 1991), there has been little research on how instructions ought to change as a person performs the procedure repeatedly (Ganier, 2004; Pea, 2004). A novice would likely require instructions that are different from those required by a person who has gained some basic proficiency with the procedure (Schnotz, 2002). The present research tested assumptions about differences in cognitive processing facilitated by linguistic and graphical representations. I posit that representations could transition from one type to another to facilitate efficient initial performance *and* subsequent learning. Furthermore, affordances designed into physical objects can provide clues about the necessary procedures. These affordances likely influence the manner in which a person uses instructions and the knowledge they gain from the instructions. This study investigated how people cognitively process various types of instructions for different types of physical objects.

Performance and Learning Outcomes

A person can either rely on instructions as an aid to perform an unfamiliar procedure or gain information from the instructions resulting in the learned ability to perform the procedure without the instructions. Performance is defined as the observed speed and accuracy with which a person can perform a procedure while using instructions, whereas learning is defined as the same observed variables without the aid of the instructions. The design of the instructions can facilitate either performance or learning (Eiriksdottir & Catrambone, 2011; Hmelo & Guzdial, 1996; Kissane, Kalyuga, Chandler, & Sweller, 2008). Specific instructions, those that indicate precise actions to take, often lead to better performance while general instructions, those that provide information requiring interpretation, often lead to better learning. But what about when people use instructions over and over and their knowledge level about the procedure changes? For procedures that are meant to be performed repeatedly, such as those involved in assembly or maintenance of objects, I posited that instructions ought to change as the procedure is repeated and the person gains knowledge. There is, however, a dearth of theoretical research on the topic (Pea, 2004; Puntambekar & Hübscher, 2005). Investigated in this study was the idea that manipulating the specificity and generality of the instructions would lead a person from efficient initial performance to robust subsequent learning.

One way to manipulate the specificity and generality of instructions is through the representation of the information. From the procedural skill literature there is a trend for graphics to be beneficial for performance and text to be beneficial for learning (Ganier, 2001; Palmiter, Elkerton, & Baggett, 1991; Watson, Butterfield, Curran, & Craig, 2010;

Yuviler-Gavish, Yechiam, & Kallai, 2011). Graphical representations readily convey spatially oriented information (Hegarty, Kriz, & Cate, 2003; Larkin & Simon, 1987). When participants performed a procedure for the first time, graphical representations resulted in better performance than simultaneous textual and graphical (Heiser, Phan, Agrawala, Tversky, & Hanrahan, 2004; Rodriguez, 2002). The graphical nature of the instructions likely aligned with the physical nature of the objects, resulting in specific information being readily conveyed and enacted, while the text demanded additional processing and interpretation before action could take place. Participants who received both types of instructions could have ignored the text, but the differences in performance provide evidence that they were paying at least some attention to the text as well as the graphics. The inclusion of linguistic aspects appears to improve learning of subsequent iterations of a procedure. Simultaneous textual and graphical representations were found to decrease initial performance but resulted in better learning (Brunyé et al., 2008), and requiring people to restate linguistic instructions in their own words resulted in better learning (Hard, Lozano, & Tversky, 2006). In each of the aforementioned findings, people were performing a physically oriented procedure, such as assembling a printer or piece of furniture. In these cases, participants learned using the same external representation for each build.

By changing the representation from build to build, instructions can be used to manipulate the person's cognitive load and processing (Ainsworth, 2006). Although it is possible for a representation to be considered task inappropriate and thus might interfere with learning (Schnotz & Bannert, 2003), I posit that transitioning representations from a predominantly graphical representation (task appropriate) to a predominantly textual representation (task inappropriate) is a way to encourage deeper processing of instructions across repetitions of a procedure. This approach will likely avoid overwhelming the person with cognitively demanding instructions for the first build and also avoid having the person mindlessly follow cognitively undemanding instructions for subsequent builds.

Underlying Cognitive Processing

Instructions must contain the information necessary for completing the procedure, and this information must be comprehendible. Although there are many models of processing procedural instructions, few specify exactly what type of information should be included. One model generally suggests using positive statements (as opposed to stating what the person should *not* do) and avoiding the need to make inferences (Bovair & Kieras, 1991). Alternatively, another model suggests that there are three types of information necessary for developing "complete" instructions (Bieger & Glock, 1984). To be considered complete, each instructional step should include the specific action to be performed, spatial information about where to perform the action (including location, orientation, or composition of an object), and organizing information about the context of the current action in relation to those actions that preceded or will follow it. It is possible to convey this information through different representations, such as a text or images, and the manner in which the mind processes these representations can vary (Kintsch, 2008). Discussed below are some models of how researchers believe people process procedural instructions of different representations. Although there are some references to graphics in these models, none of the models explicitly take into account the different processing necessary for text and graphics (Ganier, 2012; McNamara & Magliano, 2009).

Regardless, I make some generalizations about the different types of processing that likely occur when people use text versus graphics.

A textual representation requires that the person *translate* the procedural information into behavioral performance (Guthrie, Bennett, & Weber, 1991). In order to execute textual instructions, three basic steps occur. Long-term memory must be accessed to form rules, determine what rules currently exist, and update existing rules with the new information (Bovair & Kieras, 1991). Text is believed to activate more general semantic information than images (Durso & Johnson, 1979), which could be useful for learning but not performance as the additional semantic activation is likely unnecessary to perform a well-specified procedure.

Graphics processing stands in contrast to text processing. Graphics are useful for displaying location-related information (Larkin & Simon, 1987), a benefit of particular note with the spatial procedures of interest here. Graphics can contain a great deal of spatial information that is readily processed. In fact, people tend to underestimate the amount of information that can be conveyed using graphics, resulting in lower processing of the graphics (Schnotz, 2002). Implicit here is that graphics can efficiently convey a great deal of information that does not require as much cognitive processing as text (Larkin & Simon, 1987). Although it is possible to convey the necessary information using text, the instructions must be translated into an actionable representation. In fact, people tend to misremember textual instructions as pictorial instructions when learning procedures involving physical objects (Brunyé, Taylor, Rapp, & Spiro, 2006). Based on these models, additional support is provided for the notion that images provide specific information that is easily actionable while text provides general information requiring translation and the activation of additional information that is not necessarily crucial to performing the procedure.

Physical Objects as Memory Cues

The procedures in this study focus on the manipulation of physical objects. As defined by Romiszowski (1999), these procedures "depend on the recall of a possibly complex, but essentially algorithmic, procedure and the execution of a series of linked actions in sequence" (p. 464). A person perceives a stimulus in the workspace, activates relevant knowledge associated with the particular stimulus, and then performs the necessary action or actions. These procedures go beyond motor movements that are highly automatizable and discrete in nature (e.g., Adams, 1987), such as typing or rotary pursuit tasks. To operationalize the procedures of interest in the proposed research, they must include: (1) a specific end goal, not an array of possible outcomes, (2) manual manipulation of physical objects, not just bodily movements or gestures, (3) an algorithmic procedure of multiple manipulations, not just a single movement, and (4) a procedure intended for repeated execution.

Inherent to the procedures and objects of interest in this research are affordances. Affordances are physical aspects of the objects themselves that provide clues about required actions (Chemero, 2003; Jones, 2003; Turvey, 1992). For example, certain parts of the object might connect only to other pieces with similar connection interfaces, indicating to a person that this part must go in this location in a certain orientation. Engineering psychologists and those working in related fields strive to design objects that facilitate intuitive interactions (Norman, 2002; Wertsch, 1998). To the extent that not all information regarding the procedure can be made obvious to all people, instructions will be required to direct the person to perform the appropriate procedure.

The physical nature of the objects could offer the person memory cues regarding the procedure. When building an unfamiliar object for the first time, instructions can provide this information by indicating which part to select and how to place it. When building the same object for a second time, I posit that the object itself begins to take on instructive properties. When the person views the object in its current state, the shape of the missing piece might be apparent, facilitating the person in recalling which part to select. Once selected, the part might fit on the object in only one particular orientation. If decisions must be made, such as how exactly to insert the part, the physical aspects of the object could provide cues to memories formed during the first assembly. I suspected this phenomenon of the object – or completed assembly – cueing the person's memory to be dependent on the object itself. In the case where each part is unique and can be placed in only one location in one specific orientation on the object, this object would be more instructive than an object whose parts can be combined in many configurations. An object with many interchangeable parts would provide fewer memory cues as to the assembly order, although the parts of this object might be easier to place since the interchangeable nature of the parts means the actions to connect the parts to the object are be same. In this case, there are few specialized actions or orientations that must be learned. These hypothetical objects described here contain differing levels of 'knowledge in the world' (Gray & Fu, 2001). Although people are known to ignore knowledge in the world in favor of knowledge in the head because of the ease with which it can be accessed, physical objects likely represent a special case. In order to assemble the objects successfully, people are forced to confront the knowledge in the world while they attempt to place the parts. This knowledge is necessarily accessed while attempting to complete the procedure, and thus would be used to cue knowledge.

Hypotheses

The following hypotheses below address the two primary scenarios of instructional use, performing a procedure *with* and *without* the aid of the instructions. The primary manipulation in this study is the external representation (subsequently referred to as simply "representation") of the instructions, either images or text. Also manipulated is the representation of the instructions the second time the procedure is performed, again images or text. The hypotheses also take into account different types of objects with which instructions might be used, such as those with high and low affordances. Here, affordances are operationalized as physical characteristics of the objects that provide information about the assembly process, either what parts come next in the sequence or how a selected part should be oriented and placed. Affordances could affect how a person would use the instructions (particularly when performing the procedure repeatedly) and what information a person would remember about the procedure.

Initial Instructional Use

Initial instructional use is operationalized as the first time a person performs a procedure with the aid of the instructions. It was expected that those who used images would spend less time viewing the instructions and report lower cognitive workload because images require less cognitive processing whereas text necessitates the activation of more related knowledge. There were no expectations that time spent acting (or physically manipulating the parts) would be different for images and text; once a plan of action has been decided upon, either through instruction from images or text, the approach to building was not expected to be different. The aforementioned pattern was expected for both high and low affordance objects. Additionally, higher affordance objects were expected to result in more instructional views. Although a higher affordance object contained clues about which parts could or could not be placed in a certain location, the uniqueness of each part necessitated that the proper orientation and other part-specific details must be understood. Parts of higher affordance objects required that the person learned how each part was to be oriented and placed, and these orientations and placements were different for most parts. Low affordance objects, on the other hand, had many parts that were interchangeable and could be placed in multiple locations. Here, the actions to place each part were more likely to be the same, meaning a person must learn fewer precise orientation and placement actions.

Change in Instructional Use

Change in instructional use was operationalized as the differences in how instructions were used from the first to the second build. It was expected that those people who used the same instructions for both builds would demonstrate a greater reduction in instructional viewing times and reported cognitive workload. Seeing the same instructions for a second time would be easier to process as the knowledge associated with the particular representation would be at higher levels of activation because of the previous exposure. Viewing a different representation the second time would require activation of additional knowledge. Viewing text and then images would yield the greatest reductions in instructional viewing times and reported workload given that text demanded cognitive processes to translate the linguistic representation to a spatial one, and that images demanded relatively very little processing (Larkin & Simon, 1987). These differences were expected to be attenuated for low affordance objects. The interchangeable nature of the parts meant that the necessary actions to place the parts were similar, requiring overall less new processing from the first to second build. Greater reductions in instructional viewing times were expected for the higher affordance objects. The affordances were expected to provide cues as to the necessary part selection and required action. By viewing the object itself, a person would have a more distinct memory trace for the high affordance objects, thus needing to view the instructions less. Low affordance objects provided few cues about the required actions, thus a person would rely more on the instructions from the first to the second build.

Memory Build Performance

In many operationalized contexts, a desired outcome is that the person retains memory of the procedure, enabling the person to perform the procedure without the aid of the instructions. Text was expected to result in more accurate builds for low affordance objects. Although low affordance objects provided few cues about the procedure, the additional activation of semantically related knowledge (Durso & Johnson, 1979) would likely lead to more cognitive connections, thus enabling more accurate recall. For high affordance objects, it was expected that there would be no difference in the accuracy of the builds, and, in fact, that there would be few final errors. Given the lack of interchangeability of the high affordance objects, there were fewer chances to complete the assembly with errors. That said, it was expected that text would lead to less time spent correcting errors, for example, removing parts to complete a previously forgotten step.

Text would lead to a greater activation of semantically related knowledge and thus likely lead to better memory for easy to overlook steps.

Memory for Procedural Steps

Separate from acquiring the skill to assemble the physical object is to remember aspects of the procedure that are separate from the tangible aspects of building. Possessing a robust knowledge of the procedure, particularly the order in which steps must be performed, likely underpins a person's ability to solve transfer-type problems. Such problems might include disassembly of the object to access a malfunctioned part requiring replacement or understanding how the object works as a function of how the parts are arranged. An assessment tested the participants' ability to accurately judge the order of the parts used in the procedure. It was expected that having both types of instructions, images and text, would yield more accurate memory judgments for the order of the parts. A greater exposure to text, using it for both builds rather than just one or no builds, would enable more accurate judgments of part order when those parts are presented as text descriptions. The same was not expected for exposure to images. All participants built the objects, and through building with the physical parts, participants were exposed to the spatial and visual nature of the parts. Although it was possible to translate a spatial memory of the instructions to a linguistic one, this translation could result in errors. Finally, it was expected that people were more likely to report using image instructions – when they actually used text – than reporting text (when they actually used images) (Brunyé et al., 2008; Brunyé et al., 2006). This finding would confirm that text was more likely to be processed and remembered in a spatial format, mistakenly leading people to report using images.

CHAPTER 2

METHOD

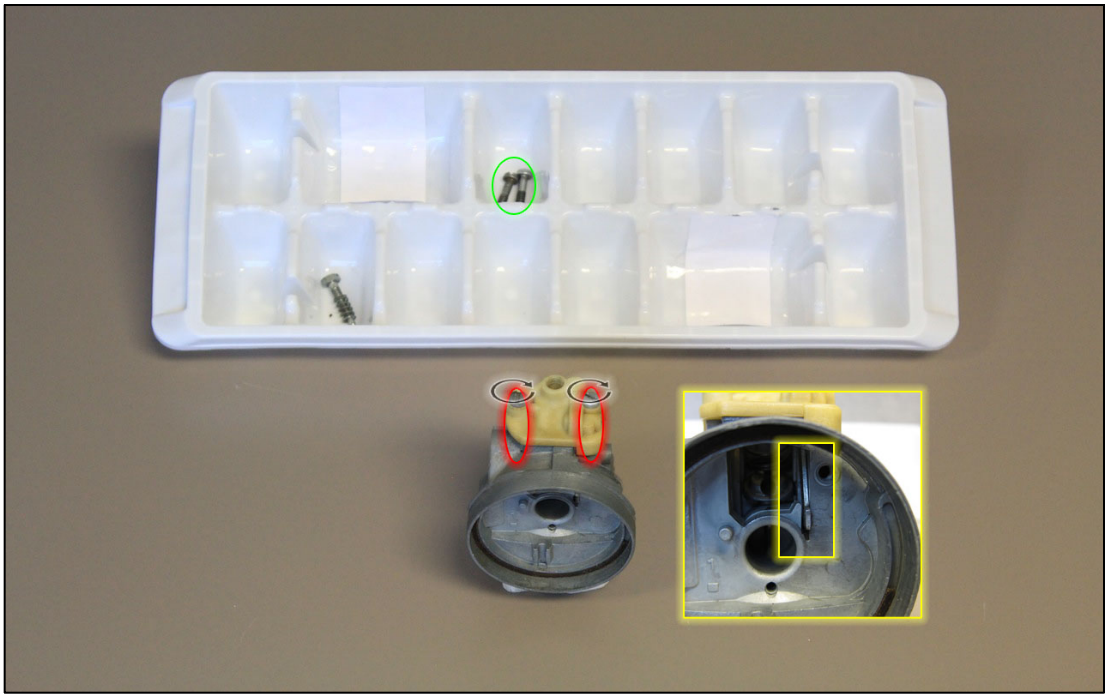
Participants

Seventy-two participants were recruited from the Georgia Tech undergraduate population via Sona. Participants had to be at least 18 years old, as the focus of this experiment was adult cognition. There were no other exclusion criteria. Participants ranged in age from 18 to 25 years, comprised 37.5% females, and reported a mean GPA of 3.46 (SD = 0.44). Participants reported either SAT or ACT individual test scores, which were converted to SAT composite scores, with a mean score of 2035 out of 2400 (SD = 150). Participants reported a mean number of earned credits of 54.4 (SD = 30.1).

Materials

Instructions

Two instructional representations were used to convey the procedural information: images and text (see Figure 1 for examples). They were designed to be informationally equivalent (Larkin & Simon, 1987) and complete (Bieger & Glock, 1984). Action verbs (i.e., place, slide, screw, and orient) for text instructions and the equivalent action arrows for image instructions, along with textual part descriptions, were developed and tested among five participants. This procedure helped to ensure that the descriptions and images unambiguously conveyed the intended meaning.



Screw the two screws with threads
up half the shafts through
the left and right sides of the yellow part.
Note: Looking into the large opening
on the front of the body, ensure
the tab fits into the slot.

Figure 1: Example instructions for one step in the assembly of the carburetor. The top half of the figure shows the image representation (notes the screws circles in the tray, two small arrows and circles showing where to place them, and the square callout box showing an important detail) and the bottom shows the text representation. Each instructional representation conveys which part to select, where and how to place the parts, and emphatic information about an easy to overlook details.

According to Bieger and Glock (1984), complete procedural instructions must include contextual, operational, and spatial information. The instructions used in this study also included descriptive information about the parts. Participants were not expected to know the names of any parts prior to the experiment, so all parts were described according to their shape and color. The text instructions comprised an operational verb, a description of the part according to color and shape, and spatial information describing the location and orientation where the part that was to be placed. Some steps included emphatic information, in which crucial details were made more salient. Contextual information was presented prior to the start of the assembly that provided the participants with a general idea of the shape and function of the object they were about to assemble. Image instructions were then developed based on the information conveyed by the text instructions.

The instructions were tested with three new pilot participants who used the instructions to build the objects while thinking aloud. The feedback gleaned informed refinement of the instructions, which were then tested in the same manner with 10 additional pilot participants. During these sessions, the researcher made notes of particular steps where participants struggled with either understanding the instructions or enacting the desired step with the physical objects. Some steps were split into two lower-level steps. For example, all pilot participants had a difficult time understanding the first step for one of the objects in which they needed to precisely orient a part while placing it in a complex housing. The researcher split this step so the participants first oriented the part outside of the housing, and then placed the properly oriented part into the housing.

Objects

All participants assembled three objects: a sweater shaver, a carburetor, and a Lego tower (see Figure 2). Some aspects of the objects were similar, each comprising ten parts and each were small enough to be easily manipulated. The objects differed according to their physical affordances indicative of their assembly procedure. The individual parts of the shaver and carburetor were unique in their color and shape, whereas some Lego parts shared either shape or color – but never both – with other parts. Additionally, the manner in which parts of the shaver and carburetor connected made it is possible to determine where a part was to be placed by looking for its respective mating surface.

Figure 2: The three objects assembled by participants: a sweater shaver, a moped carburetor, and a Lego tower.

Part Order Verification Assessment

To probe participant's memory, a part order verification assessment was developed based on Brunyé et al. (2008). The participant was presented with two parts at a time. The participant was instructed to determine if the parts were displayed in the same order that they were used in the assembly procedure. For example, if the third part from the assembly was displayed on the left and the fifth part on the right, the participant was instructed to indicate that these parts were displayed in the same order by typing ‘Y’. The participant responded to 20 of these pairs, half presented as images of the parts and half presented as text descriptions of the parts. For each of these two groups of stimuli, half were presented in the correct order and half were presented in the incorrect order. The computer recorded the accuracy of each determination for each pair.

Procedure

All participants assembled in order: the shaver, the carburetor, and the Lego tower. To assemble a given object, participants received either images for both assemblies (Image—Image), text for both (Text—Text), images for the first and text for the second (Image—Text), or text and then images (Text—Image). Participants were assigned to different instructional conditions for each object. For example, a participant might receive Image—Image instructions for the shaver, Image—Text for the carburetor, and Text—Text for the tower. Counterbalancing the four instructional conditions yielded 24 possible permutations. Using an effect size of 0.40 for build times when using image compared to text instructions (Brunyé et al., 2008; Hochmitz & Yuviler-Gavish, 2011) and a power of 0.80, it was estimated that 68 participants were necessary. Repeating the 24 instructional sequences thrice yielded a total of 72 participants.

There order of the objects was not counterbalanced for two reasons. First, there were few expected carryover effects. Although the participants might have had a clearer idea of the procedure after assembling the shaver, there was little, or no, information in the instructions for each of the three objects that would provide a benefit for assembling a subsequent object. Had the experiment been performed with objects that were similar with systematic variations (e.g., three Lego objects, one with three different color bricks, one with five, and one with seven), counterbalancing would be necessary because assembling one of those Lego objects could provide a benefit for another Lego object. Second, 144 participants would have been needed to counterbalance all instructional conditions *and* all orders of object assembly, more than twice as many participants as needed.

Participants were first assessed on their perceptual speed, ideational fluency, and spatial orientation (Ekstrom, French, & Harman, 1979). The specific tests administered were the Identical Pictures Test, the Thing Categories Test, and the Cube Comparison Test, respectively. The experimenter then provided the participant with an overview of the workspace and a tutorial of the instructional experience (see Figure 3 for an example of the workspace). Sitting closest to the participant on the desk was a tablet computer (Google Nexus 7, 2012 version). Directly behind the tablet sat a video camera (Logitech HD Portable 1080p Webcam C615 with Autofocus) that displayed a live video stream on the experimenter's computer. The object to be assembled sat just behind the video camera, and all of the parts from which to select sat right behind that.

Figure 3: An overview of the workspace. A tablet computer sat in front of objects to be assembled by the participant. Instructions were displayed on the tablet in the form of images or text.

Already setup in the workspace was a tutorial item (a Coleman portable battery-powered lantern) that was partially disassembled. The experimenter sat next to the participant during the tutorial. The experimenter described that the participant must place both thumbs on the tablet to view the instructions. The participant was encouraged to view the instructions for as long as necessary before performing the instructed action. After viewing the first step, the participant performed the action (i.e., inserting a battery). Once the battery was inserted, the participant was instructed to press DONE on the tablet. The experimenter selected CORRECT on his computer and described how this feedback is displayed on the tablet for a short time before the next step begins. After completing the second step, the experimenter selected INCORRECT to demonstrate the incorrect feedback screen. The experimenter showed two other steps, with all four steps together covering the range of actions a participant would perform in the assemblies: placing, sliding, screwing, and orienting. The experimenter then repeated the same four steps, this time with text. The experimenter invited the participant to ask questions about the instructions and workspace. A detailed mockup of the experimenter and participant interfaces can be found in the Appendix.

The experimenter then discussed the concept of workload with the participant. The participant read the definition of the six sources of workload from the NASA TLX (Hart & Staveland, 1988). The participant then completed the calibration procedure in which the participant selected which source of workload might contribute more to workload in a pairwise presentation of the six sources. The experimenter then showed the participant a blank workload rating sheet.

The participant then assembled the shaver using image or text instructions according to their assigned condition. The instructional system logged all actions on the tablet, feedback, and any notes regarding errors that the experimenter wrote. The participant then reported their workload on each of the six NASA TLX scales. The experimenter then asked the question, “Do you have any comments or feedback about the instructions and how they helped you assemble the object?” The experimenter elicited comments about how the participant *used* the instructions as opposed to their preference for the instructions. The aforementioned procedure was repeated (i.e., assembly with instructions, workload rating, interview question) for the second time with the same object. After the two builds, an additional question was asked, “How, if at all, did you use the instructions differently from the first to the second build?” Then, the participant completed the part order verification assessment and reported what type of instructions they used earlier in the experiment, images or text, for the two assemblies. Finally, the participant assembled the shaver from memory while video camera recorded the process. This procedure was then repeated for the carburetor and for the Lego tower.

After the three assemblies and assessments, the experimenter asked if the participant had any final comments or feedback about the instructions and the objects. The participant was instructed to return to the lab eight to ten days later for a follow-up session. At this session, the participant repeated the following for each of the three objects: responded to the part order verification assessment, reported what type of instructions they used during the previous assemblies, and assembled the object from memory. An overview of the procedure is shown in Figure 4. These assessments were performed in the same order as when participants learned with instructions (first the shaver, and then carburetor, and then the Legos).

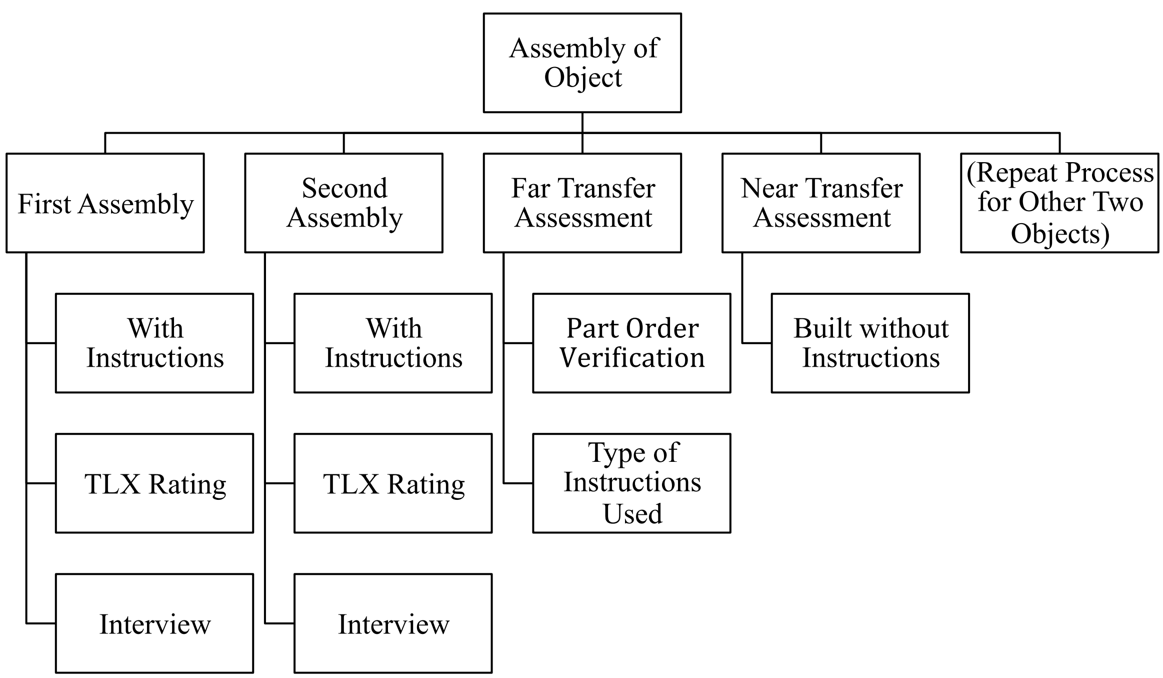


Figure 4: A flow chart of the main part of the experiment. Prior to assembling the objects, participants completed some initial tasks (i.e., consent form, background questionnaire,

and abilities tests) as well as a tutorial (i.e., description of the tablet, workspace, process for viewing instructions, and receiving feedback).

Data Coding

Abilities Assessments

The three abilities assessments were coded according to the developers' instructions (Ekstrom et al., 1979). The Identical Pictures Test, which measured the participants' perceptual speed, was scored as the number correctly marked pictures minus the number of incorrectly marked pictures. The Thing Categories Test, which measured the participants' ideal fluency, was scored as the number of things that belong to the given categories. The Cube Comparisons Test, which measured the participants' spatial aibilty, was scored as the number correctly marked cubes minus the number of incorrectly marked cubes. Each test comprised two parts, resulting in two scores for each test. These two scores for each of the three respective tests were averaged yielding three scores for each participant.

Instructional Usage

While using the instructions, the system automatically logged each interaction with the instructions on the part of the participant and the experimenter. The system generated a timestamp when: the participant started viewing the instructions, the participant stopped viewing the instructions, the participant tapped DONE, and the experimenter provided feedback (CORRECT, INCORRECT, or PAUSE (meaning a participant made the same error twice and the experimenter will demonstrate the step)). The system also recorded notes written by the experimenter describing any errors made. The durations between each timestamp were parsed into one of two time variables summed across all steps for one assembly of one object: time spent viewing the instructions (viewing time) and time spent assembling the object (acting time). If the time between starting the step for the second time (in the case of making an error trying again) and the first instructional view was greater than 10 seconds, this duration was added to the acting time total, otherwise it was added to the viewing time total. For any acting times that were less than three seconds in duration, these times were added to the viewing time total. In this case, participants either 'flickered' the instructions (i.e., tapped repeatedly switch between views with and without the instructions) or they started to reach for parts but decided to return quickly back to the instructions. The cutoffs of 10 and 3 seconds were determined through observation of the participants while using the instructions. Furthermore, the total number of instructional views, number of corrected errors, and number of mediated errors were recorded.

Reported Workload

The overall workload scores were calculated according to the developers' procedures (Hart & Staveland, 1988). For the initial calibration task, a tally was performed to determine the weighting factor for each scale. The six individual ratings of workload, from 0 to 100 in increments of 5, were multiplied by their respective weighting factors, summed, and divided by 15 to determine the overall workload score. Six overall workload scores were computed for each participant, one for each of the two builds for each of the three objects.

Interview

The interviews were coded for themes. *A priori* codes were used to code the interviews. Information in the interview data not captured by the codes was used to develop post *hoc codes*. Both sets of codes were combined to form the final list of codes (see **Error! Reference source not found.**). Two independent raters coded the transcripts.

Table 1:

Definitions of the codes used for the interview statements.

First Build	Participant Descriptions
Passive	clear, straightforward, no ambiguity; no comments; gist, skimmed
Active	attention, read carefully, re-read, looked for key features/locations, clarification
Constructive	figured it out, convert, interpret, translate, unclear, confusion
Second Build	
Passive	clear, straightforward, no ambiguity; no comments; gist, skimmed
Active	attention, read carefully, re-read, looked for key features/locations, clarification
Constructive	figured it out, convert, interpret, translate, unclear, confusion
Difference in Processing	
Decreased	decreased reliance, used prior knowledge, verifying steps, mostly for order
Constant	used in the same way, no difference, first helped to understand second, clarifying
Increased	increased reliance on the instructions, gaining different knowledge

Part Order and Instruction Type Memory Assessment

The part order assessments were coded as either correct or incorrect for each of the 20 assessment items. Six scores were calculated for each participant: overall speed and accuracy, speed and accuracy for the pairs presented as images, and speed and accuracy for the pairs presented as text. The memory type assessment was coded as correct, image instructions mistakenly remembered as text, or text instructions mistakenly remembered as images.

Assembly from Memory

The video recordings of the three memory assemblies were coded for time spent thinking and time spent acting. Act time was operationalized as the time a part was in motion, whereas thinking time was when no parts were moving. Three coders coded four different videos and checked each other's work. Differences were discussed among the three raters. The three raters then each coded one-third of the videos. One rater checked and made corrections to all videos to ensure consistency. The coding was performed using InqScribe software.

Final errors, errors still present after the participant finished, were coded. Final errors were present only with the Lego towers. The final towers were coded for correct color placement, correct shape placement, and overall part locations. Correct color placement was coded by looking at all possible 16 locations in the tower (the four levels of the four corners of the tower) and assigning one point for each correct color, regardless if it was the correct block. Correct shape placement was coded by assigning one point for each shape in the correct location even if it was the wrong color. Finally, part location was coded for each of the 10 parts. Each part received up to three points for being in the correct location. One point was assigned if the block was on the correct side (left or right), one if it was in the proper depth (front or back), and at the right height in the tower.

CHAPTER 3

RESULTS: INSTRUCTIONAL USE

The results presented in this section describe how participants used the instructions. There are two primary subsections of results: how participants used instructions for the first build of each object and how instructional use changed from the first to the second build of each object. Each subsection contains a series of analyses on a variety of dependent variables. Each of the two subsections concludes with a summary of the results and initial interpretations about the patterns of results.

Initial Instructional Use

This subsection addresses how participants used the instructions during the first build of each object. The independent variable was the type of instructions used by participants, either image or text. The dependent variables were instructional viewing time, number of instructional views, reported mental workload, acting time, and build accuracy.

Instructional Viewing Time

For each of the three objects, mean instructional viewing times were compared between those who used image and text instructions for the first build. For each of the three objects, image instructions consistently resulted in significantly shorter viewing times (see Table 2). There was a significant interaction between instruction type and object type ($F(2,210) = 13.35$, $p = .001$, $MSE = 31045$, $\eta_p^2 = .113$) suggesting that the magnitude of the differences between image and text instructions were different for each of the three objects.

Table 2:

Analysis of viewing time by instructional condition for the first build of each object.

	Instructions					
Object	Image	Text	*F*(1,68)	*p*	*MSE*	η_p^2
Shaver	125.9 (41.0)	221.1 (72.4)	48.83	.001	163349	.418
Carburetor	118.4 (42.8)	226.7 (69.4)	63.25	.001	211331	.482
Lego	38.1 (9.4)	68.9 (17.5)	84.67	.001	17049	.555

Note: Means have the units of seconds, and standard deviations are in parentheses. Critical $\alpha = 0.05$.

Reported Mental Workload

For each of the three objects, mean reported mental workloads were compared between those who used image and text instructions for the first build. For the shaver and carburetor, image instructions resulted in significantly lower reported workload, whereas no difference was found for Lego (see Table 3). There was not a significant interaction between instruction type and object type ($F(2,210) = 1.59$, $p = .207$, $MSE = 9974$, $\eta_p^2 = .015$), therefore interpretations about differences between the instruction types must be made with caution.

Table 3:

Analysis of reported mental workload by instructional condition for the first build of each object.

Object	Instructions					
	Image	Text	$F(1,68)$	p	MSE	η_p^2
Shaver	87.9 (88.9)	131.8 (92.5)	4.12	.046	34672	.057
Carburetor	100.6 (80.1)	143.5 (90.3)	4.45	.039	33153	.061
Lego	34.4 (64.6)	37.1 (49.8)	.04	.847	125	.001

Note: Analyses were performed on the weighted ratings of the mental workload subscale. Means can range from 0 (low reported workload) to 500 (high reported workload), and standard deviations are in parentheses. Critical $\alpha = 0.05$.

Number of Instructional Views

For each of the three objects, mean number of instructional views were compared between those who used image and text instructions for the first build. For the shaver and carburetor, image instructions resulted in significantly *greater* number of views, whereas no difference was found for Lego (see Table 4). There was a significant interaction between instruction type and object type ($F(2,210) = 4.50$, $p = .012$, $MSE = 154$, $\eta_p^2 = .041$) suggesting that the magnitude of the differences between image and text instructions were different for each of the three objects.

Table 4:

Analysis of the number of instructional views by instructional condition for the first build of each object.

	Instructions					
Object	Image	Text	*F*(1,68)	*p*	*MSE*	η_p^2
Shaver	25.7 (6.6)	22.3 (4.3)	6.75	.011	210	.090
Carburetor	29.3 (9.1)	22.3 (6.2)	14.20	.001	882	.173
Lego	14.2 (4.1)	13.0 (2.4)	2.34	.131	26	.033

Note: Means have the units of number of views, and standard deviations are in parentheses. Critical α = 0.05.

Acting Time

For each of the three objects, mean acting times were compared between those who used image and text instructions for the first build. For the shaver and carburetor, image instructions resulted in significantly shorter acting times, whereas no difference was found for Lego (see Table 5). There was a significant interaction between instruction type and object type ($F(2,210) = 3.56$, $p = .030$, $MSE = 41458$, $\eta_p^2 = .033$) suggesting that the magnitude of the differences between image and text instructions were different for each of the three objects.

Table 5:

Analysis of acting time by instructional condition for the first build of each object.

	Instructions					
Object	Image	Text	*F*(1,68)	*p*	*MSE*	η_p^2
Shaver	220.6 (62.2)	306.3 (105)	17.30	.001	132362	.203
Carburetor	312.0 (111)	393.8 (204)	4.54	.037	120394	.063
Lego	51.5 (25.9)	52.2 (12.1)	.02	.881	9	.000

Note: Means have the units of seconds, and standard deviations are in parentheses. Critical α = 0.05.

Build Accuracy

For each step in the first build with instructions, the numbers of corrected and mediated errors were measured. A 'corrected error' was defined as when the participant made a mistake yet was able to fix the mistake given a second attempt and view of the instructions. A 'mediated error' was defined as when the participant made a mistake on a build step but was not able to fix the mistake on the second attempt. In this case, the experimenter demonstrated the step. Overall error rates were low for each of the three objects. For all steps performed by all participants, corrected error rates were 6.8%, 8.7%, and 4.7% for the shaver, carburetor, and Legos respectively. Mediated error rates were 2.2%, 3.7%, and 0.3% respectively.

There were some instances where text resulted in significantly more corrected and mediated errors. Across all the steps of the shaver, 16 participants who used text instructions experienced an average of 1.3 mediated errors, whereas only one participant who used images experienced one mediated error. For the Legos, 26 participants who used text experienced an average of 1.2 corrected errors across the 10 total steps of the assembly, whereas three participants who used images each experienced one corrected error. In all other cases (i.e., corrected errors for the shaver and carburetor and mediated errors for the Legos), the number of corrected and mediated errors did not differ significantly between images and text (see Table 6).

Table 6:

Analysis of number of corrected and mediated errors for the first build of each object.

	Corrected Errors				Mediated Errors			
Object	Image Rank	Text Rank	*Z*	*p*	Image Rank	Text Rank	*Z*	*p*
Shaver	38.42	34.58	-.849	.396	28.96	44.04	-4.134	.001
Carburetor	40.99	32.01	-1.934	.053	32.50	40.50	-1.999	.046
Lego	24.79	48.21	-5.482	.001	35.50	37.50	-1.424	.154

Note: Mann-Whitney U test. A 'corrected error' was defined as when the participant made a mistake on a build step yet was able to fix the mistake given a second attempt and view of the instructions. A 'mediated error' was defined as when the participant made a mistake on a build step but was not able to fix the mistake on the second attempt. In this case, the experimenter demonstrated the step. Critical $\alpha = 0.05$.

Summary of Initial Instructional Use

The pattern of significant results differed among the three objects (see Table 7). Shorter instructional viewing times were found for those who used images for all three objects. Although the interaction was not significant, there was a trend for lower reported mental workload for those who used images for the shaver and carburetor. When reporting workload, participants were instructed to focus on their use of the instructions. These results suggest that images generally required less time to process for the shaver and carburetor.

For number of instructional views, images yielded *more* than text for the shaver and carburetor. When considered with shorter acting times for the same two objects, these results suggest a pattern of use where the participant spent a short amount of time acting and returned to the instructions more frequently when using images. With text instructions, participants spent more time figuring out how to place the complex parts of the shaver and carburetor, but they did not need to return to the text instructions as frequently.

In regards to errors, the shaver and carburetor did not yield a difference in corrected errors, but there were more mediated errors for text instructions. The opposite was true for Legos. These results suggested that it was easy for participants to fix their misunderstandings about the textual Lego instructions, whereas participants were more likely to experience errors they could not correct on their own with the shaver and carburetor when using text instructions.

Table 7:

Summary of the dependent measures for initial instructional use for each object.

Object	Viewing Time	Mental Workload	Number of Views	Acting Time	Corrected Errors	Mediated Errors
Shaver	I < T (.42)	I < T (.06)	I > T (.09)	I < T (.20)	=	I < T
Carburetor	I < T (.48)	I < T (.06)	I > T (.17)	I < T (.06)	=	I < T
Lego	I < T (.56)	= (.00)	= (.03)	= (.00)	I < T	=

Note: ‘I’ stands for image instructions and ‘T’ stands for text. Effect sizes (η_p^2) are in parentheses. The units for viewing and acting time are seconds. The units for number of views, corrected errors, and mediated errors are the number of occurrences. The scale for mental workload ranges from 0 to 500 and is unit-less. An ‘=‘ indicates there was not a significant difference between image and text instructions.

Changes in Instructional Use

This subsection addresses how instructional use changed from the first to the second build of each object. The independent variables was the type of instructions received for the first build and the type received for the second build. The dependent variables were changes (from the first to second build with instructions) in instructional viewing time, changes in number of instructional views, changes in reported mental workload, changes in acting time, and changes in build accuracy.

Change in Instructional Viewing Time

For each of the three objects, the changes from the first to the second build with instructions in viewing times by instructional condition were analyzed. Omnibus tests revealed there was a significant overall reduction in viewing times for each of the three objects. Shaver: Build 1, $M = 173.5$ ($SD = 75.6$); Build 2, $M = 66.5$ ($SD = 24.6$) ($F(1,68) = 277.41$, $p = .001$, $MSE = 412196$, $\eta_p^2 = .803$). Carburetor: Build 1, $M = 172.5$ ($SD = 79.1$); Build 2, $M = 68.0$ ($SD = 34.3$) ($F(1,68) = 260.88$, $p = .001$, $MSE = 393894$, $\eta_p^2 = .793$). Lego: Build 1, $M = 53.5$ ($SD = 20.8$); Build 2, $M = 47.6$ ($SD = 17.5$) ($F(1,68) = 10.10$, $p = .002$, $MSE = 1251$, $\eta_p^2 = .129$). There was a significant interaction between instruction type and object type ($F(6,204) = 3.85$, $p = .001$, $MSE = 6184$, $\eta_p^2 = .102$) suggesting that the magnitude of the differences between build 1 and 2 were different for each of the three objects.

For the shaver, all conditions resulted in significant decreases. For the carburetor, all but the Image—Text condition resulted in significant decreases. For the Lego, all but the Image—Image condition resulted in significant decreases (see Table 8).

Table 8:

Analysis of the change, from the first to the second build with instructions, in viewing time by instructional condition for each object.

	Build w/Instructions				
Object	First	Second	*F*(1,68)	*p*	η_p^2
Shaver					
Image—Image	114.0 (26.0)	50.8 (7.8)	24.22	.001	.263
Image—Text	137.7 (49.8)	95.3 (25.9)	10.90	.002	.138
Text—Image	237.5 (76.7)	54.4 (14.6)	203.10	.001	.749
Text—Text	204.8 (65.8)	65.5 (17.0)	117.43	.001	.633
Carburetor					
Image—Image	121.5 (42.5)	43.6 (6.5)	36.16	.001	.347
Image—Text	115.2 (44.2)	108.7 (35.7)	.26	.615	.004
Text—Image	214.9 (78.4)	50.7 (14.3)	160.79	.001	.703
Text—Text	238.5 (58.9)	68.8 (26.4)	171.73	.001	.716
Lego					
Image—Image	36.7 (8.9)	38.2 (9.4)	.156	.694	.002
Image—Text	39.5 (9.9)	63.4 (19.8)	41.39	.001	.378
Text—Image	68.5 (18.7)	36.2 (7.3)	75.97	.001	.528
Text—Text	69.3 (69.2)	52.7 (14.8)	19.95	.001	.227

Note: Means have the units of seconds, and standard deviations are in parentheses. Critical $\alpha = 0.013$ using a Bonferroni correction for four simultaneous analyses.

Change in Reported Mental Workload

For each of the three objects, the changes from the first to the second build with instructions in reported mental workloads by instructional condition were analyzed. Omnibus tests revealed there was a significant overall reduction in reported mental workload for each of the three objects. Shaver: Build 1, $M = 109.9$ ($SD = 92.8$); Build 2, $M = 51.8$ ($SD = 64.9$) ($F(1,68) = 46.72$, $p = .001$, $MSE = 121336$, $\eta_p^2 = .407$). Carburetor: Build 1, $M = 122.0$ ($SD = 87.4$); Build 2, $M = 57.7$ ($SD = 59.7$) ($F(1,68) = 77.12$, $p = .001$, $MSE = 148867$, $\eta_p^2 = .531$). Lego: Build 1, $M = 35.8$ ($SD = 57.3$); Build 2, $M = 29.4$ ($SD = 54.6$) ($F(1,68) = 4.69$, $p = .034$, $MSE = 1469$, $\eta_p^2 = .065$). There was not a significant interaction between instruction type and object type ($F(6,204) = 0.57$, p = .754, $MSE = 4697$, $\eta_p^2 = .016$), therefore interpretations about differences in the changes from build 1 to 2 and across the objects must be made with caution.

For the shaver, those who received text for the first build reported significant decreases. For the carburetor, all but those who used Image—Text reported significant decreases. For the Lego, only those who received one of each type (i.e., Image—Text and Text—Image) reported significant decreases (see Table 9).

Table 9:

Analysis of the change, from the first to the second build with instructions, in reported mental workload by instructional condition for each object.

	Build w/Instructions				
Object	First	Second	$F(1,68)$	p	η_p^2
Shaver					
Image—Image	84.4 (92.6)	41.9 (64.1)	6.26	.015	.084
Image—Text	91.4 (87.6)	78.3 (80.5)	.59	.445	.009
Text—Image	141.9 (99.6)	47.2 (58.1)	31.09	.000	.314
Text—Text	121.7 (86.5)	39.7 (49.5)	23.27	.000	.255
Carburetor					
Image—Image	92.8 (77.4)	47.8 (54.9)	9.44	.003	.122
Image—Text	108.3 (84.2)	74.7 (53.3)	5.27	.025	.072
Text—Image	139.2 (82.7)	66.7 (68.9)	24.51	.001	.265
Text—Text	147.8 (99.5)	41.7 (59.2)	52.50	.001	.436
Lego					
Image—Image	29.4 (45.5)	17.8 (34.1)	3.91	.052	.054
Image—Text	39.4 (80.4)	57.8 (79.4)	9.66	.003	.124
Text—Image	26.7 (29.5)	7.2 (17.0)	10.86	.002	.138
Text—Text	47.5 (63.4)	34.7 (56.2)	4.69	.034	.065

Note: Analyses were performed on the weighted ratings of the mental workload subscale. Means can range from 0 (low reported workload) to 500 (high reported workload), and standard deviations are in parentheses. Critical $\alpha = 0.013$ using a Bonferroni correction for four simultaneous analyses.

Change in Number of Instructional Views

For each of the three objects, the changes from the first to the second build with instructions in number of instructional views by instructional condition were analyzed. Omnibus tests revealed there was a significant overall reduction in number of views for the shaver and carburetor but not the Lego. Shaver: Build 1, $M = 24.0$ ($SD = 5.8$); Build 2, $M = 17.3$ ($SD = 3.6$) ($F(1,68) = 160.4$, $p = .001$, $MSE = 1613$, $\eta_p^2 = .702$). Carburetor: Build 1, $M = 25.8$ ($SD = 8.5$); Build 2, $M = 14.8$ ($SD = 2.9$) ($F(1,68) = 200.46$, $p = .001$, $MSE = 4334$, $\eta_p^2 = .747$). Lego: Build 1, $M = 13.6$ ($SD = 3.4$); Build 2, $M = 13.4$ ($SD = 2.9$) ($F(1,68) = .66$, $p = .419$, $MSE = 2$, $\eta_p^2 = .010$). There was not a significant interaction between instruction type and object type ($F(6,204) = 1.278$, $p = .269$, $MSE = 41$, $\eta_p^2 = .036$), therefore interpretations about differences in the changes from build 1 to 2 and across the objects must be made with caution.

For the shaver and the carburetor, all conditions resulted in significant decreases. For the Lego, only the Image—Text conditioned resulted in a significant decrease (see Table 10).

Table 10:

Analysis of the change, from the first to the second build with instructions, in the number of instructional views by instructional condition for each object.

	Build w/Instructions				
Object	First	Second	$F(1,68)$	p	η_p^2
Shaver					
Image—Image	25.8 (7.0)	20.0 (5.5)	29.88	.001	.305
Image—Text	25.7 (6.2)	15.9 (2.3)	84.59	.001	.554
Text—Image	23.1 (4.4)	17.7 (2.5)	26.53	.001	.281
Text—Text	21.5 (4.1)	15.7 (1.5)	30.45	.001	.309
Carburetor					
Image—Image	28.4 (8.1)	15.3 (2.5)	70.95	.001	.511
Image—Text	30.1 (10.3)	14.4 (2.9)	102.17	.001	.600
Text—Image	21.6 (7.3)	15.2 (3.5)	16.99	.001	.200
Text—Text	22.9 (4.9)	14.1 (2.7)	32.07	.001	.321
Lego					
Image—Image	13.2 (2.9)	14.0 (3.0)	1.65	.203	.024
Image—Text	15.3 (4.9)	13.6 (3.2)	6.60	.012	.088
Text—Image	13.6 (3.1)	13.8 (3.0)	.07	.798	.001
Text—Text	12.4 (1.4)	12.1 (1.9)	.36	.551	.005

Note: Means have the units of number of views, and standard deviations are in parentheses. Critical α = 0.013 using a Bonferroni correction for four simultaneous analyses.

Change in Acting Time

For each of the three objects, the changes from the first to the second build with instructions in acting times by instructional condition were analyzed. For the shaver and carburetor but not the Lego, omnibus tests revealed there was a significant overall reduction in acting times. Shaver: Build 1, M = 263.5 (SD = 96.1); Build 2, M = 157.0 (SD = 33.7) ($F(1,68)$ = 109.5, p = .001, MSE = 407706, η_p^2 = .617). Carburetor: Build 1, M = 352.9 (SD = 168.3); Build 2, M = 202.5 (SD = 63.2) ($F(1,68)$ = 90.57, p = .001, MSE = 814935, η_p^2 = .571). Lego: Build 1, M = 51.9 (SD = 20.1); Build 2, M = 52.3 (SD = 34.2) ($F(1,68)$ = .02, p = .889, MSE = 5, η_p^2 = .000).There was a significant interaction between instruction type and object type ($F(6,204) = 2.14$, $p = .050$, $MSE = 19761$, $\eta_p^2 = .059$) suggesting that the magnitude of the differences between build 1 and 2 were different for each of the three objects.

For the shaver and the carburetor, all conditions resulted in significantly faster acting times, whereas none of the conditions resulted in significant decreases for the Lego (see Table 11).

Table 11:

Analysis of the change, from the first to the second build with instructions, in acting time by instructional condition for each object.

	Build w/Instructions				
Object	First	Second	$F(1,68)$	p	η_p^2
Shaver					
Image—Image	210.8 (46.2)	150.2 (24.2)	8.87	.004	.115
Image—Text	230.4 (75.0)	171.7 (42.0)	8.33	.005	.109
Text—Image	306.8 (92.7)	161.6 (31.3)	51.02	.001	.429
Text—Text	305.8 (119)	144.7 (31.0)	62.79	.001	.480
Carburetor					
Image—Image	324.6 (110)	201.5 (55.1)	15.17	.001	.182
Image—Text	299.5 (114)	217.3 (62.9)	6.75	.011	.090
Text—Image	346.2 (119)	174.8 (30.9)	29.37	.001	.302
Text—Text	441.5 (258)	216.3 (86.6)	50.70	.001	.427
Lego					
Image—Image	47.3 (22.0)	60.1 (58.6)	5.93	.017	.080
Image—Text	55.7 (29.3)	60.0 (28.4)	.67	.417	.010
Text—Image	53.7 (8.9)	45.0 (11.7)	2.70	.105	.038
Text—Text	50.8 (14.8)	43.9 (15.8)	1.76	.189	.025

Note: Means have the units of seconds, and standard deviations are in parentheses. Critical $\alpha = 0.013$ using a Bonferroni correction for four simultaneous analyses.

Change in Build Accuracy

For each step in the second build with instructions, the numbers of corrected and mediated errors were measured. A ‘corrected error’ was defined as when the participant made a mistake on a build step yet was able to fix the mistake given a second attempt and view of the instructions. A ‘mediated error’ was defined as when the participant made a mistake on a build step but was not able to fix the mistake on the second attempt. In this case, the experimenter demonstrated the step. Overall error rates were low for each of the three objects during the second build. For all steps performed across all participants, corrected error rates were 0.9%, 2.1%, and 2.5% for the shaver, carburetor, and Legos respectively. Mediated error rates were 0.0%, 0.2%, and 0.4% respectively.

In terms of corrected errors, all participants in all condition made fewer errors during the second build except for: Image—Text and Text—Text with the carburetor and Image—Image and Image—Text for the Legos. In terms of mediated errors, there were no reductions in the number of errors except for: Image—Text for the shaver, and Text—Image and Text—Text for the carburetor For a summary of the changes in errors, see Table 12.

Table 12:

Analysis of the change, from the first to the second build with instructions, in the number of corrected and mediated errors by instructional condition for each object.

	Corrected Errors			Mediated Errors		
Object	*t*	Ranks	*p*	*t*	Ranks	*p*
Shaver						
Image—Image	-2.970	10/0/8	.003	0.000	0/0/18	1.000
Image—Text	-2.730	13/1/4	.006	-1.000	1/0/17	.317
Text—Image	-2.949	12/1/5	.003	-3.274	12/0/6	.001
Text—Text	-3.071	11/0/7	.002	-1.890	4/0/14	.059
Carburetor						
Image—Image	-3.334	15/1/2	.001	-.816	2/1/15	.414
Image—Text	-1.941	12/4/2	.052	-1.730	5/1/12	.084
Text—Image	-2.972	10/0/8	.003	-2.530	7/0/11	.011
Text—Text	-2.360	11/3/4	.018	-2.565	8/0/10	.010
Lego						
Image—Image	0.000	1/1/16	1.000	0.000	0/0/18	1.000
Image—Text	-2.456	0/7/11	.014	-1.732	0/3/15	.083
Text—Image	-3.500	13/0/5	.001	-1.000	1/0/17	.317
Text—Text	-2.546	11/2/5	.011	-1.000	1/0/17	.317

Note: Wilcoxon signed ranks tests. All ranks are build 2 minus build 1. The ranks column shows negative/positive/tie ranks. A 'corrected error' was defined as when the participant made a mistake on a build step yet was able to fix the mistake given a second attempt and view of the instructions. A 'mediated error' was defined as when the participant made a mistake on a build step but was not able to fix the mistake on the second attempt. In this case, the experimenter demonstrated the step. Critical $\alpha = 0.013$.

Summary of Changes in Instructional Use

The pattern of significant results differed among the three objects (see Table 13). Two general conclusions arose from the pattern of results. The first was that those participants who used images and then text appeared to maintain more consistent or longer processing times across the two builds. Considering viewing time for the carburetor, for this condition only was there no decrease in this measure. For the shaver while using Image—Text, there was no decrease in viewing time (which was also true for the Image—Image condition). For the Legos while using Image—Text, there was an *increase* in this measure (despite no change or decreases for all other conditions). Taken together, these results suggested that images followed by text could help to maintain a more consistent level of cognitive processing.

As evidenced by the change in acting time, much of the learning occurred during the first build, particularly for the shaver and the carburetor. With all types of instructions for the shaver and carburetor, acting time decreased from the first to the second build. This was not the case for the Legos. With all types of instructions for the Legos, there was no significant decrease in acting time. The Legos comprised simple, repetitive steps, whereas the shaver and carburetor comprised relatively more complex and idiosyncratic actions. These results suggested that much of the precise motor actions are learned during the first build, likely through a combination of the information presented in the instructions and through trial-and-error with the physical parts.

Table 13:

Summary of the dependent measures for the changes in instructional use for each object.

Object	Viewing Time	Mental Workload	Number of Views	Acting Time	Corrected Errors	Mediated Errors
Shaver						
Image—Image	– (.26)	= (.08)	– (.31)	– (.12)	–	=
Image—Text	– (.14)	= (.01)	– (.55)	– (.11)	–	=
Text—Image	– (.75)	– (.31)	– (.28)	– (.43)	–	–
Text—Text	– (.63)	– (.26)	– (.31)	– (.48)	–	=
Carburetor						
Image—Image	– (.35)	– (.12)	– (.51)	– (.18)	–	=
Image—Text	= (.00)	= (.07)	– (.60)	– (.09)	=	=
Text—Image	– (.70)	– (.27)	– (.20)	– (.30)	–	–
Text—Text	– (.72)	– (.44)	– (.32)	– (.43)	=	–
Lego						
Image—Image	= (.00)	= (.05)	= (.02)	= (.08)	=	=
Image—Text	+ (.38)	+ (.12)	– (.09)	= (.01)	=	=
Text—Image	– (.53)	– (.14)	= (.00)	= (.04)	–	=
Text—Text	– (.23)	= (.07)	= (.01)	= (.03)	–	=

Note: The units for viewing and acting time are seconds. Effect sizes (η_p^2) are in parentheses. The units for views, corrected errors, and mediated errors are the number of occurrences. The scale for mental workload ranges from 0 to 500 and is unit-less. An '=' indicates there was not a significant difference between build 1 and build 2.

CHAPTER 4

RESULTS: LEARNING OUTCOMES

The results presented in this section describe the outcomes subsequent to interacting with the instructions. There are two primary subsections of results: near transfer assessments in which participants built the objects without instructions and far transfer (Barnett & Ceci, 2002) assessments in which participants made judgments from memory about the order of the procedure. Each of the two subsections concludes with a summary of the results and initial interpretations about the patterns of results.

Near Transfer Outcomes

This subsection addresses how near transfer outcomes changed between the immediate and delayed assessments. The independent variables were the type of instructions used for the two builds. The dependent variables were changes (from the immediate assessments right after using the instructions to delayed assessments 8-10 days later) in thinking time, changes in acting time, and changes in Lego build accuracy. Tests of *a priori* hypotheses were followed by regressions that investigated other potential variables influencing differences in outcomes.

Change in Thinking Time

For each of the three objects, the changes from the immediate to delayed memory build assessment in thinking times by instructional condition were analyzed. Omnibus tests revealed there was a significant overall increase in thinking times for only the carburetor. Shaver: Immediate assessment, M = 21.5 (SD = 9.0); Delayed, M = 23.3 (SD = 12.1) (F(1,64) = 2.13, p = .149, MSE = 116, η_p^2 = .032). Carburetor: Immediate assessment, M = 20.9 (SD = 16.9); Delayed, M = 34.3 (SD = 33.2) (F(1,63) = 9.32, p = .003, MSE = 6003, η_p^2 = .129). Lego: Immediate assessment, M = 62.3 (SD = 47.9); Delayed, M = 52.6 (SD = 58.4) (F(1,63) = 1.86, p = .178, MSE = 3538, η_p^2 = .029).

Despite the lack of omnibus significance for the shaver and Lego, simple main effects were still analyzed. There were no instructional conditions for any of the objects that resulted in a significant change in thinking times between the two assessments (see Table 14).

Table 14:

Analysis of the change, from the immediate to the delay assessment, in thinking times by instructional condition.

	Assessment				
Object	Immediate	Delayed	F	p	η_p^2
Shaver			df = (1,64)		
Image—Image	20.1 (6.5)	19.4 (5.8)	.08	.780	.001
Image—Text	22.1 (7.0)	26.5 (16)	2.82	.098	.042
Text—Image	22.4 (9.3)	26.9 (14)	3.23	.077	.048
Text—Text	21.4 (12)	20.6 (9.3)	.11	.740	.002
Carburetor			df = (1,63)		
Image—Image	23.8 (18)	39.1 (32)	2.92	.093	.044
Image—Text	17.8 (7.4)	31.7 (18)	2.57	.114	.039
Text—Image	17.2 (6.3)	25.0 (13)	.81	.372	.013
Text—Text	25.2 (27)	41.6 (54)	3.57	.063	.054
Lego			df = (1,63)		
Image—Image	58.8 (51)	52.9 (53)	.15	.704	.002
Image—Text	69.0 (51)	58.2 (49)	.55	.461	.009
Text—Image	53.9 (37)	56.2 (56)	.03	.876	.000
Text—Text	68.3 (55)	41.4 (41)	2.84	.097	.043

Note: Means have the units of seconds, and standard deviations are in parentheses. Critical α = 0.013 using a Bonferroni correction for four simultaneous analyses.

Change in Acting Time

For each of the three objects, the changes from the immediate to delayed memory build assessment in acting times by instructional condition were analyzed. Omnibus tests revealed there was a significant overall increase in acting times for the shaver and carburetor but not the Lego. Shaver: Immediate assessment, $M = 92.5$ ($SD = 26.1$); Delayed, M = 100.9 ($SD = 37.8$) ($F(1,64) = 4.23$, $p = .044$, $MSE = 2405$, $\eta_p^2 = .062$). Carburetor: Immediate assessment, $M = 135.4$ ($SD = 36.1$); Delayed, M = 187.9 ($SD = 97.6$) ($F(1,63) = 20.68$, $p = .001$, $MSE = 92027$, $\eta_p^2 = .247$). Lego: Immediate assessment, $M = 69.1$ ($SD = 51.1$); Delayed, M = 61.5 ($SD = 49.5$) ($F(1,64) = 1.08$, $p = .303$, $MSE = 2128$, $\eta_p^2 = .017$).

Despite the lack of significance for the Lego, simple main effects were still analyzed. Only for the Text—Text condition for the carburetor did analyses yield a significant increase in acting times between the two assessments (see Table 15).

Table 15:

Analysis of the change, from the immediate to the delay assessment, in acting times by instructional condition.

	Assessment				
Object	Immediate	Delayed	F	p	η_p^2
Shaver			df = (1,64)		
Image—Image	97.3 (32.3)	94.9 (18.8)	.08	.774	.001
Image—Text	91.1 (19.1)	109 (41.5)	4.45	.039	.065
Text—Image	97.2 (22.0)	99.6 (42.9)	.09	.760	.001
Text—Text	84.9 (26.6)	101 (44.2)	3.92	.052	.058
Carburetor			df = (1,63)		
Image—Image	144 (38.3)	196 (68.9)	4.77	.033	.070
Image—Text	129 (27.5)	172 (78.5)	3.52	.065	.053
Text—Image	124 (30.9)	161 (53.0)	2.59	.112	.040
Text—Text	145 (44.1)	223 (153)	11.76	.001	.157
Lego			df = (1,64)		
Image—Image	80.5 (63.1)	58.4 (46.5)	2.12	.150	.032
Image—Text	63.0 (33.0)	60.1 (43.0)	.04	.847	.001
Text—Image	62.4 (40.6)	66.5 (72.1)	.08	.783	.001
Text—Text	71.4 (66.0)	60.6 (25.8)	.44	.509	.007

Note: Means have the units of seconds, and standard deviations are in parentheses. Critical $\alpha = 0.013$ using a Bonferroni correction for four simultaneous analyses.

Change in Lego Build Accuracy

Given the nature of the objects, only the Lego towers could result in final errors. Three scores were coded: placing a block of the correct color regardless of the shape of the block (‘Color’); placing a block of the correct shape regardless of the color of the block (‘Shape’); and the location of each block compared ideal location in three dimensions (‘Location’). These scores ranged from 0 (completely incorrect) to 1 (completely correct). All three scores were used in a multivariate analysis to determine the overall change in accuracy from the immediate to delayed assessment (see Table 16). Although all instruction types resulted in decreased accuracy at the time of delayed assessment, those who used images only demonstrated the smallest effect size whereas those who used text only demonstrated the largest effect size.

Table 16:

Analysis of the change, from the immediate to the delay assessment, in Lego accuracy by instructional condition.

	Assessment				
Object	Immediate	Delayed	$F(1,67)$	p	η_p^2
Image—Image	.77 (.04)	.60 (.06)	8.95	.004	.118
Image—Text	.73 (.04)	.55 (.06)	9.99	.002	.130
Text—Image	.76 (.04)	.57 (.06)	12.11	.001	.153
Text—Text	.82 (.04)	.57 (.06)	17.59	.001	.208

Note: Means are the proportion of correct outcomes, and standard errors are in parentheses.

Other Potential Near Transfer Predictors

Considering the limited number of significant differences in terms of performance by instructional condition, regression analyses were performed to investigate other potential variables influencing performance. The following variables were used in a stepwise regression: spatial ability, perceptual speed, ideational fluency, overall reported mental workload, overall instructional viewing time, overall acting time with instructions, and overall number of instructional views. Some variables entered the regression equations for thinking time (see Table 17) and acting time (see Table 18). The most notable finding is that acting time (when using the instructions) most consistently predicted thinking and acting times (when building without instructions).

Table 17:

Analysis of potential predictors, for both the immediate and delayed assessment, for thinking times for each object.

Object	*Beta*	*t* (df)	*p*	R^2	*F(1)*	*p*
Shaver, Immediate						
Acting time	.27	2.35 (63)	.022	.073	5.52	.022
Shaver, Delayed						
Acting time	.49	4.52 (59)	.001	.236	20.43	.001
Carb, Immediate						
(no predictors)						
Carb, Delayed						
Acting time	.59	5.99 (59)	.001	.352	35.88	.001
Lego, Immediate						
(no predictors)						
Lego, Delayed						
(no predictors)						

Note: Six separate regressions were performed (3 objects x 2 DV assessments). Predictors were: spatial ability, perceptual speed, ideational fluency, overall reported mental workload, overall instructional viewing time, overall acting time with instructions, and overall number of instructional views. Stepwise regressions were performed with Pin = .05 and Pout = .10.

Table 18:

Analysis of potential predictors, for both the immediate and delayed assessment, for acting times for each object.

Object	*Beta*	*t* (df)	*p*	R^2	*F(*df*)*	*p*
Shaver, Immediate						
Acting time	.43	4.09 (63)	.001	.157	13.04 (1)	.001
Mental workload	-.26	-2.57 (62)	.012	.215	9.44 (2)	.001
Idea fluency	.25	2.43 (61)	.018	.265	8.17 (3)	.001
Number views	.22	2.08 (60)	.042	.309	7.50 (4)	.001
Shaver, Delayed						
Acting time	.59	5.89 (59)	.001	.345	34.73 (1)	.001
Carb, Immediate						
Acting time	.31	2.68 (62)	.009	.134	10.67 (1)	.002
Perceptual speed	-.23	-2.01 (61)	.049	.182	7.59 (2)	.001
Carb, Delayed						
Acting time	.59	5.98 (59)	.001	.352	35.79 (1)	.001
Lego, Immediate						
(no predictors)						
Lego, Delayed						
Spatial Ability	-.28	-2.35 (59)	.022	.077	5.51 (1)	.022

Note: Six separate regressions were performed (3 objects x 2 DV assessments). Predictors were: spatial ability, perceptual speed, ideational fluency, overall reported mental workload, overall instructional viewing time, overall acting time with instructions, and overall number of instructional views. Stepwise regressions were performed with Pin = .05 and Pout = .10.

Summary of Near Transfer Outcomes

An unexpected finding was that there were no changes in thinking or acting time from immediate to delayed assessment for any condition or object (except Text—Text for carburetor, increase in acting time). That said, there were some significant results that differed among the three objects. Of note was that acting time (while using the instructions) positively predicted thinking and acting times (while building without the instructions) for the shaver and the carburetor in seven of the eight regressions. These predictors were not significant in the Lego regressions. These results suggested that those participants who spent more time attempting to figure out what to do also spent more time retrieving procedural knowledge from memory.

The instructional conditions, however, yielded some insights into the Lego performance. Although all conditions demonstrated decreased accuracy, the Image—Image condition yielded the small decrease in accuracy over the delay, whereas Text—Text yielded the largest decrease (see Table 19). It appears that there is a benefit to using image instructions to help create more robust knowledge structures.

Table 19:

Summary of the near transfer assessment results.

	Thinking Time		Acting Time	
Object	Immediate	Delayed	Immediate	Delayed
Shaver	+Acting time	+Acting time	+Acting time -Mental workload +Idea fluency +Number views	+Acting time
Carburetor	(none)	+Acting time	+Acting time -Perceptual speed	+Acting time
Lego	(none)	(none)	(none)	-Spatial ability

Lego Instructions	Change in Accuracy
Image—Image	– (.12)
Image—Text	– (.13)
Text—Image	– (.15)
Text—Text	– (.21)

Note: The Lego object comprised bricks of various colors and shapes. Effect sizes (η_p^2) are in parentheses. Each participant build was compared to the ideal outcome.

Far Transfer Outcomes

This subsection addresses how far transfer outcomes changed between the immediate and delayed assessments. The independent variables were the type of instructions used for the two builds. The dependent variables were accuracy of procedure order judgments when presented as image and as text stimuli. Tests of *a priori* hypotheses were followed by regressions that investigated other potential variables influencing differences in outcomes.

Accuracy of Procedure Order Judgments with Images

For each of the three objects, the changes from the immediate to delayed procedure order assessment accuracy – when using *image* stimuli – by instructional condition were analyzed. There was a significant overall decrease in accuracy for all objects. Shaver: Immediate assessment, $M = .96$ ($SD = .06$); Delayed, M = .86 ($SD = .13$) ($F(1,67) = 41.94$, $p = .001$, $MSE = .3$, $\eta_p^2 = .385$). Carburetor: Immediate assessment, $M = .88$ ($SD = .12$); Delayed, M = .77 ($SD = .15$) ($F(1,67) = 37.31$, $p = .001$, $MSE = .4$, $\eta_p^2 = .358$). Lego: Immediate assessment, $M = .60$ ($SD = .15$); Delayed, M = .55 ($SD = .16$) ($F(1,67) = 8.40$, $p = .005$, $MSE = .1$, $\eta_p^2 = .111$). There was not a significant interaction between instruction type and object type ($F(6,201) = .90$, $p = .498$, $MSE = 1$, $\eta_p^2 = .026$), therefore interpretations about differences the immediate and delayed assessments and across the objects must be made with caution.

For the shaver, there were significant reductions in accuracy for all conditions except the Image—Image condition. For the carburetor, there were significant reductions in accuracy for those who first used text (i.e., Text—Image and Text—Text). For the Lego, there were significant reductions in accuracy for those who used the same instructions for both builds (i.e., Image—Image and Text—Text) (see Table 20).

Table 20:

Analysis of the change, from the immediate to the delay assessment, in accuracy of far transfer outcomes with image stimuli by instructional condition.

	Assessment				
Object	Immediate	Delayed	$F(1,67)$	p	η_p^2
Shaver					
Image—Image	.94 (.07)	.86 (.15)	6.22	.015	.085
Image—Text	.97 (.08)	.84 (.13)	18.74	.001	.219
Text—Image	.97 (.05)	.89 (.12)	6.81	.001	.092
Text—Text	.96 (.06)	.86 (.12)	12.55	.001	.158
Carburetor					
Image—Image	.85 (.14)	.76 (.16)	6.10	.016	.083
Image—Text	.84 (.13)	.75 (.17)	5.02	.028	.070
Text—Image	.93 (.08)	.82 (.10)	9.88	.002	.129
Text—Text	.92 (.09)	.77 (.12)	19.25	.001	.223
Lego					
Image—Image	.67 (.18)	.56 (.15)	9.04	.004	.119
Image—Text	.57 (.14)	.51 (.15)	2.97	.089	.042
Text—Image	.57 (.12)	.63 (.15)	3.03	.086	.043
Text—Text	.61 (.14)	.51 (.15)	8.66	.007	.103

Note: Means are the proportion of correct judgments, and standard deviations are in parentheses. Critical α = 0.013 using a Bonferroni correction for four simultaneous analyses.

Accuracy of Procedure Order Judgments with Text

For each of the three objects, the changes from the immediate to delayed procedure order assessment accuracy – when using *text* stimuli – by instructional condition were analyzed. There was a significant overall decrease in accuracy for all objects. Shaver: Immediate assessment, $M = .89$ ($SD = .13$); Delayed, M = .78 ($SD = .13$) ($F(1,67) = 45.70$, $p = .001$, $MSE = .4$, $\eta_p^2 = .406$). Carburetor: Immediate assessment, $M = .90$ ($SD = .11$); Delayed, M = .79 ($SD = .17$) ($F(1,67) = 28.34$, $p = .001$, $MSE = .4$, $\eta_p^2 = .297$). Lego: Immediate assessment, $M = .69$ ($SD = .15$); Delayed, M = .62 ($SD = .17$) ($F(1,67) = 12.46$, $p = .001$, $MSE = .2$, $\eta_p^2 = .157$). There was not a significant interaction between instruction type and object type ($F(6,201) = 0.71$, $p = .642$, $MSE = 1$, $\eta_p^2 = .021$), therefore interpretations about differences the immediate and delayed assessments and across the objects must be made with caution.

For the shaver, there were significant reductions in accuracy for all conditions except the Image—Image condition. For the carburetor, there were significant reductions in accuracy for all conditions except the Image—Image condition. For the Lego, there were no significant reductions in accuracy for any condition (see Table 21).

Table 21:

Analysis of the change, from the immediate to the delay assessment, in accuracy of far transfer outcomes with text stimuli by instructional condition.

	Assessment				
Object	Immediate	Delayed	$F(1,67)$	p	η_p^2
Shaver					
Image—Image	.85 (.16)	.78 (.17)	5.09	.027	.071
Image—Text	.88 (.12)	.76 (.11)	13.76	.001	.170
Text—Image	.90 (.12)	.77 (.09)	15.04	.001	.183
Text—Text	.94 (.11)	.82 (.16)	13.76	.001	.170
Carburetor					
Image—Image	.86 (.12)	.81 (.13)	1.12	.295	.016
Image—Text	.90 (.14)	.78 (.15)	8.14	.006	.108
Text—Image	.92 (.07)	.79 (.19)	9.22	.003	.121
Text—Text	.94 (.09)	.79 (.20)	13.67	.001	.169
Lego					
Image—Image	.71 (.16)	.63 (.18)	2.74	.102	.039
Image—Text	.67 (.18)	.61 (.14)	1.96	.166	.028
Text—Image	.73 (.11)	.63 (.15)	4.69	.034	.065
Text—Text	.67 (.15)	.59 (.21)	3.37	.071	.048

Note: Means are the proportion of correct judgments, and standard deviations are in parentheses. Critical $\alpha = 0.013$ using a Bonferroni correction for four simultaneous analyses.

Other Potential Far Transfer Predictors

Regression analyses were performed to investigate other potential variables influencing accuracy of procedure order judgments when presented with image stimuli (see Table 22) and text stimuli (see Table 23). The following variables were used in a stepwise regression: spatial ability, perceptual speed, ideational fluency, overall reported mental workload, overall instructional viewing time, overall acting time with instructions, and overall number of instructional views. There appeared to be no patterns of significant predictors in the results.

Table 22:

Analysis of potential predictors, for both the immediate and delayed assessment, of far transfer outcomes with image stimuli.

Object	*Beta*	*t* (df)	*p*	R^2	*F*(df)	*p*
Shaver, Immediate						
(no predictors)						
Shaver, Delayed						
(no predictors)						
Carb, Immediate						
Idea fluency	.26	2.39 (63)	.020	.089	6.86 (1)	.011
Number views	-.25	-2.29 (62)	.025	.148	6.00 (2)	.004
Perceptual speed	.22	2.03 (61)	.046	.197	5.56 (3)	.002
Carb, Delayed						
Spatial Ability	.40	3.58 (62)	.001	.157	12.83 (1)	.001
Lego, Immediate						
(no predictors)						
Lego, Delayed						
(no predictors)						

Note: Six separate regressions were performed (3 objects x 2 DV assessments). Predictors were: spatial ability, perceptual speed, ideational fluency, overall reported mental workload, overall instructional viewing time, overall acting time with instructions, and overall number of instructional views. Stepwise regressions were performed with Pin = .05 and Pout = .10.

Table 23:

Analysis of potential predictors, for both the immediate and delayed assessment, of far transfer outcomes with text stimuli.

Object	*Beta*	*t* (df)	*p*	R^2	*F*(df)	*p*
Shaver, Immediate						
(no predictors)						
Shaver, Delayed						
Spatial Ability	.26	2.28 (62)	.026	.070	5.17 (1)	.026
Carb, Immediate						
Acting time	-.24	-2.03 (63)	.047	.055	4.10 (1)	.047
Carb, Delayed						
Acting time	-.34	-2.99 (62)	.004	.115	8.95 (1)	.004
Lego, Immediate						
Acting time	.32	2.80 (63)	.007	.101	7.84 (1)	.007
Lego, Delayed						
(no predictors)						

Note: Six separate regressions were performed (3 objects x 2 DV assessments). Predictors were: spatial ability, perceptual speed, ideational fluency, overall reported mental workload, overall instructional viewing time, overall acting time with instructions, and overall number of instructional views. Stepwise regressions were performed with Pin = .05 and Pout = .10.

Memory for Instructional Representation

Participants were asked to remember what type of instructions – image or text – they used. In total, each participant made 12 judgments: for each of the three objects both immediately and delayed, each participant was asked which type they used for the first build and the second build. Participants were more likely to mistakenly report having used images (when they actually used text), $\chi^2(1) = 6.26$, p = 0.012. For the 12 judgments, 0.6/12 were on average reported incorrectly as images whereas 0.4/12 were on average reported incorrectly as text.

Summary of Far Transfer Outcomes.

The pattern of significant results differed among the three objects (see Table 24), although care must be taken when interpreting these results as the instruction by object interactions were not significant. It is possible that image only instructions yielded no decrease in performance for all three objects and both stimuli types. It was expected that those who received text instructions might have performed better when assessed with the text stimuli, but the results suggested that image instructions could result in more robust and transferable knowledge. Although it appeared for Legos that most instructional conditions resulted in non-significant reductions in accuracy of image or text judgments, overall performance was hardly above chance. At the delayed assessment, accuracy ranged from 59% to 63%. (In comparison, for the shaver and the carburetor, initial accuracy for both the image and text stimuli ranged from 84-97% and dropped to 76-89% after the delay, which was well above chance.) Regression analyses yielded little additional insight about what might be occurring.

Table 24:

Summary of the far transfer assessment results.

Object	Accuracy with Image Stimuli	Accuracy with Text Stimuli
Shaver		
Image—Image	= (.09)	= (.07)
Image—Text	– (.22)	– (.17)
Text—Image	– (.09)	– (.18)
Text—Text	– (.16)	– (.17)
Carburetor		
Image—Image	= (.08)	= (.02)
Image—Text	= (.07)	– (.11)
Text—Image	– (.13)	– (.12)
Text—Text	– (.22)	– (.17)
Lego		
Image—Image	– (.12)	= (.04)
Image—Text	= (.04)	= (.03)
Text—Image	= (.04)	= (.07)
Text—Text	– (.10)	= (.05)

	Image stimuli		Text stimuli	
Object	Immediate	Delayed	Immediate	Delayed
Shaver	(none)	(none)	(none)	+Spatial ability
Carburetor	+Idea fluency -Number views +Perceptual speed	+Spatial ability	-Acting time	-Acting time
Lego	(none)	(none)	+Acting time	(none)

Note: Image and text stimuli analyses investigated changes in order accuracy for each of the four conditions for each of the three objects. Effect sizes (η_p^2) are in parentheses. Predictor analyses used regressions to investigate variables that predicted image and text judgment performance for each of the three objects.

CHAPTER 5

DISCUSSION

The primary aim of this study was to better understand how people use instructions while gaining knowledge about procedural skills involving physical objects. Investigated were two sources of knowledge: instructions and the physical objects themselves. The primary manipulation of the instructions was the external representation of instructions, either images or text. These instructions were then used in the context of three different objects, two of which exhibited more affordances. This manipulation of the instructions, when applied in the context of objects possessing different characteristics, yielded patterns of results that differed among the objects. The Results section was organized in a temporal fashion, first addressing the first build with instructions followed by the second build, and then addressing the near and far transfer outcomes. The Discussion section, however, first considers the instructions and then the objects as sources of information.

Instructions as a Source of Information

As expected for the first build, images were more readily processed and acted upon, generally resulting in fewer errors. Considering the spatial nature of all three objects, it is reasonable that visiospatial instructions (i.e., images) were easier for people to understand and execute.

There was evidence that when images were followed by text, relatively more consistent levels of cognitive processing were maintained as measured by viewing time and reported mental workload. This finding suggests a simple way to maintain more consistent time spent with the instructions while performing physically oriented procedures. These more consistent levels might lead to increased task vigilance or overall increased levels of satisfaction on the part of a learner, but further research is needed.

There was some evidence that the type of instruction can influence the type of information that is acquired. Images resulted in the smallest decrease in procedural order memory accuracy (and text the largest) from the immediate to delayed assessment. Images possibly resulted in more robust knowledge, even when applied to linguistic tasks that required the person to make memory judgments about the order of the procedure. This evidence was unexpected; it was expected that those who used text would perform better at text-based tasks. Although there were some potential benefits of including text, it was possible that text introduced excessive extraneous cognitive workload, thus interfering with mental model development.

Physical Objects as Sources of Information

Despite the many differences in how participants used the instructions, the type of instructions used played only a small role in outcomes. I believe that the act of exploring, manipulating, and successfully assembling physical objects plays a large role in the knowledge gained.

Recovery from instructional misunderstandings was more likely to occur with objects comprising familiar parts and repeated actions, as was the case with Legos. In those cases where the actions to be performed were generally more unique for each step of the assembly (as with the shaver and carburetor), much of the exploration and interpretation occurred when orienting and placing the parts. This exploration was found to be more demanding with objects comprising uniquely shaped parts as evidenced by acting times and number of instructional views. Yet once the necessary actions were correctly performed on the first attempt, performance greatly improved during the second attempt.

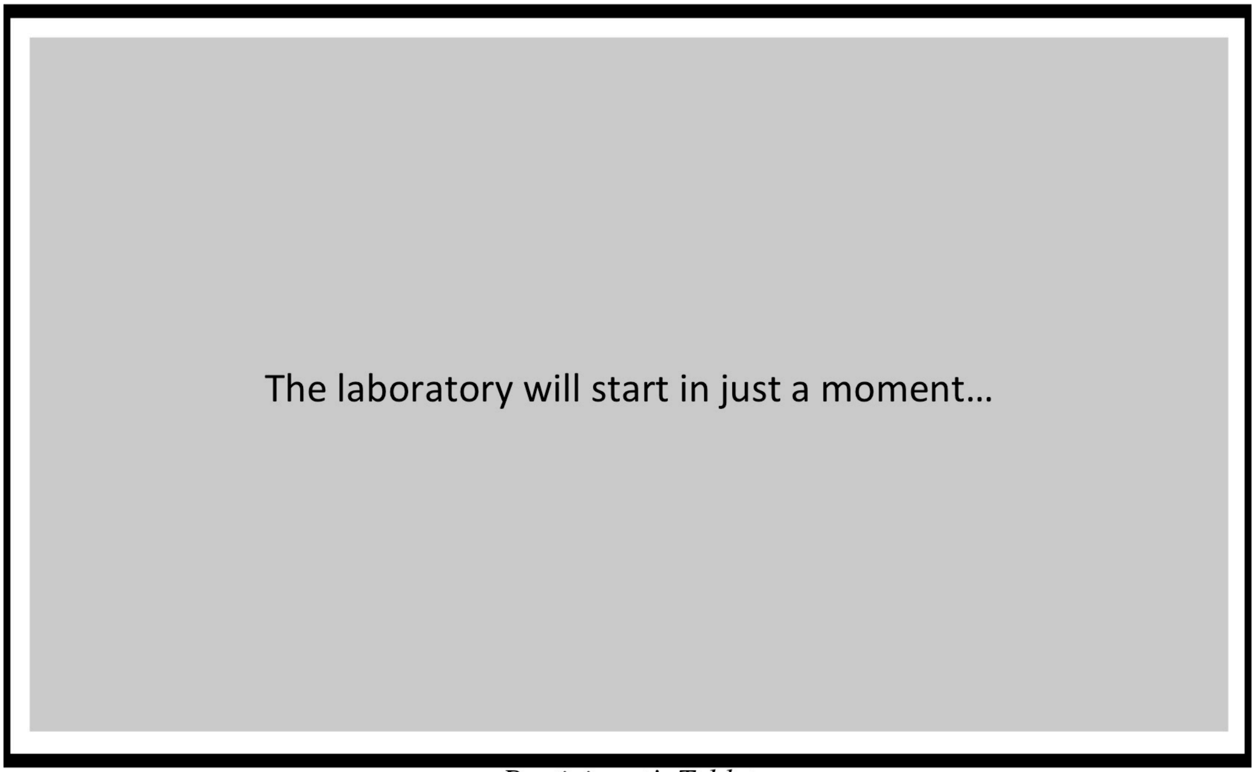
When attempting to remember the procedure, it appeared that for objects with uniquely shaped parts, people can more readily reconstruct the order of the procedure, perhaps as a result of more well organized knowledge. When the parts were less unique and had less distinguishing characteristics, it was harder for the person to mentally organize the steps in such a way to accurately recall them later.

Creating Knowledge through Integrating Information

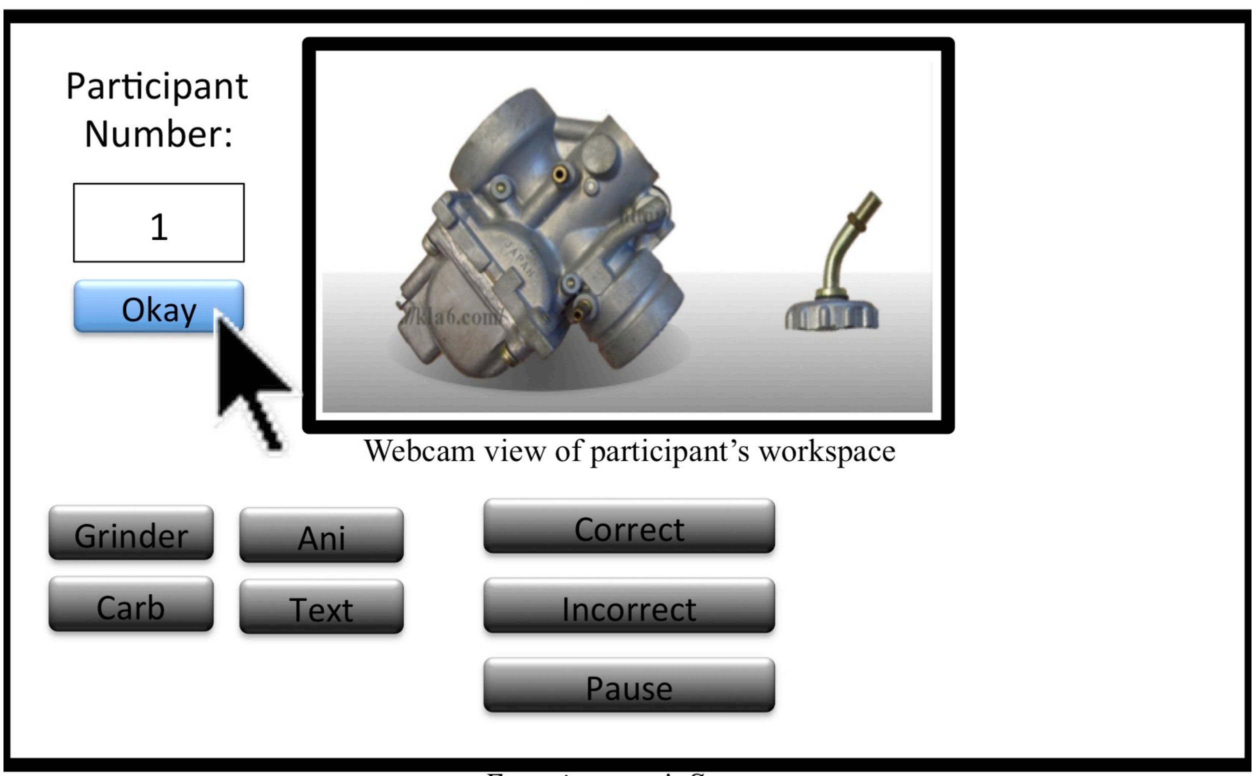
The primary finding for this study is that although people might use instructions presented as images or text very differently when interacting with physical objects, the characteristics of the objects themselves play an important role in the knowledge people acquire. More research is needed to better differentiate the role of the instructions and the role of the physical objects. One idea for a follow-up study is to have participants study only the instructions without building simultaneously, and then have them perform the near and far transfer assessments. Other participants would build the objects without instructions (but be given a model or image of the final object) and again complete the transfer assessments. These conditions would yield additional insight into which knowledge and skills are gained from the instructions and which are gained from the physical experience of building. Additionally, more research is needed in systematically understand the physical characteristics of the objects that result in outcome differences.

APPENDIX: MOCKUP OF THE INSTRUCTIONAL SYSTEM

Experimenter inputs the participant number

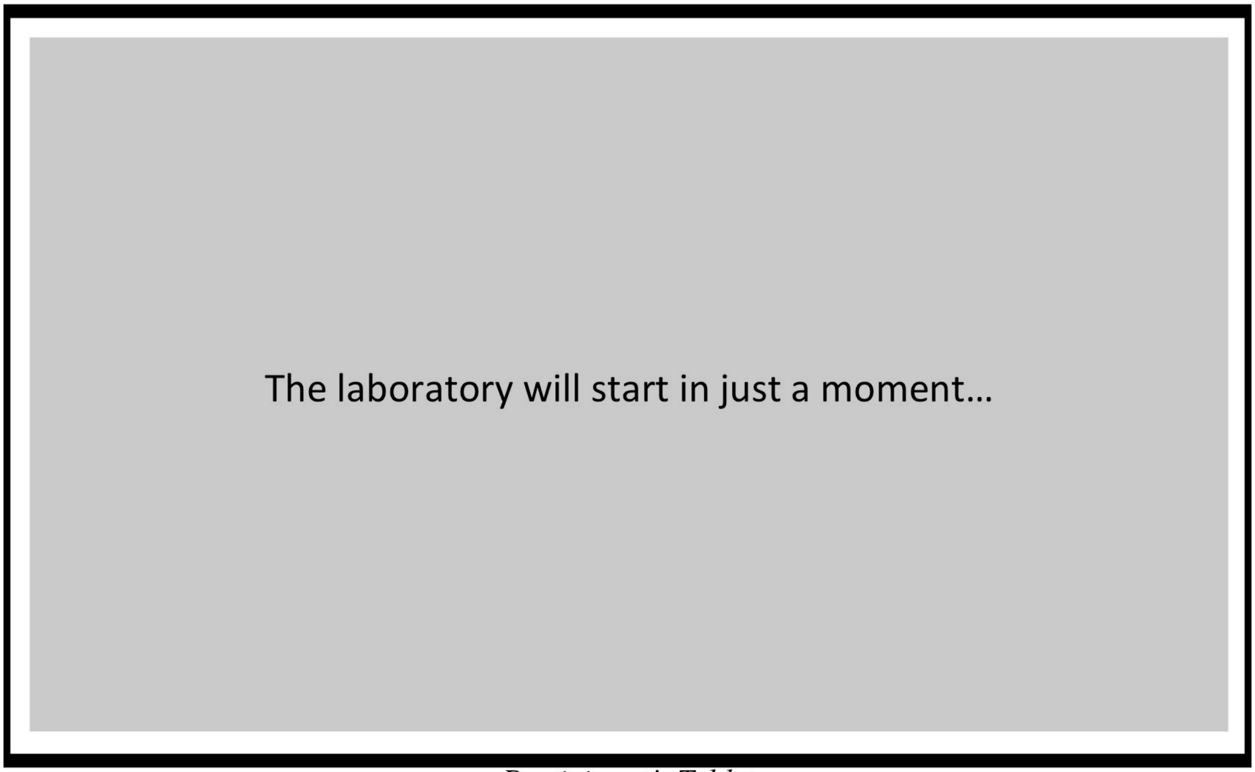


Participant's Tablet

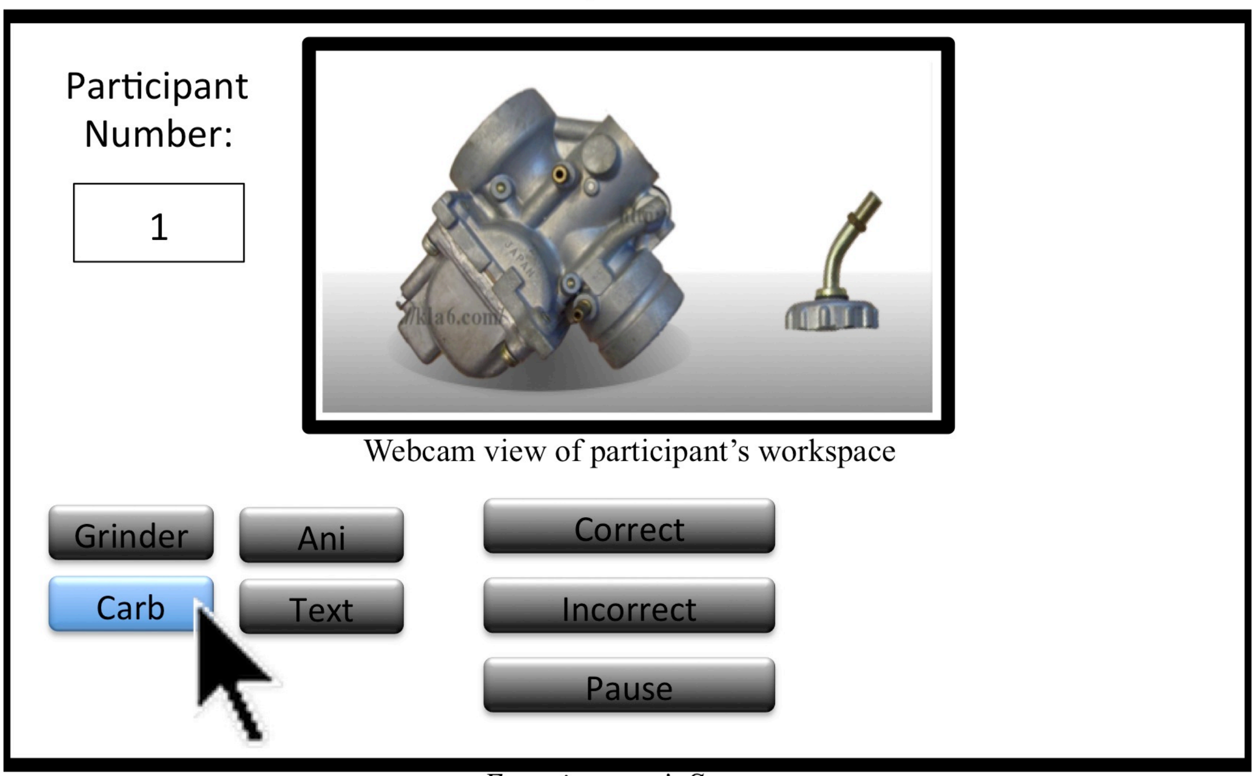


Experimenter's Screen

Experimenter selects appropriate object

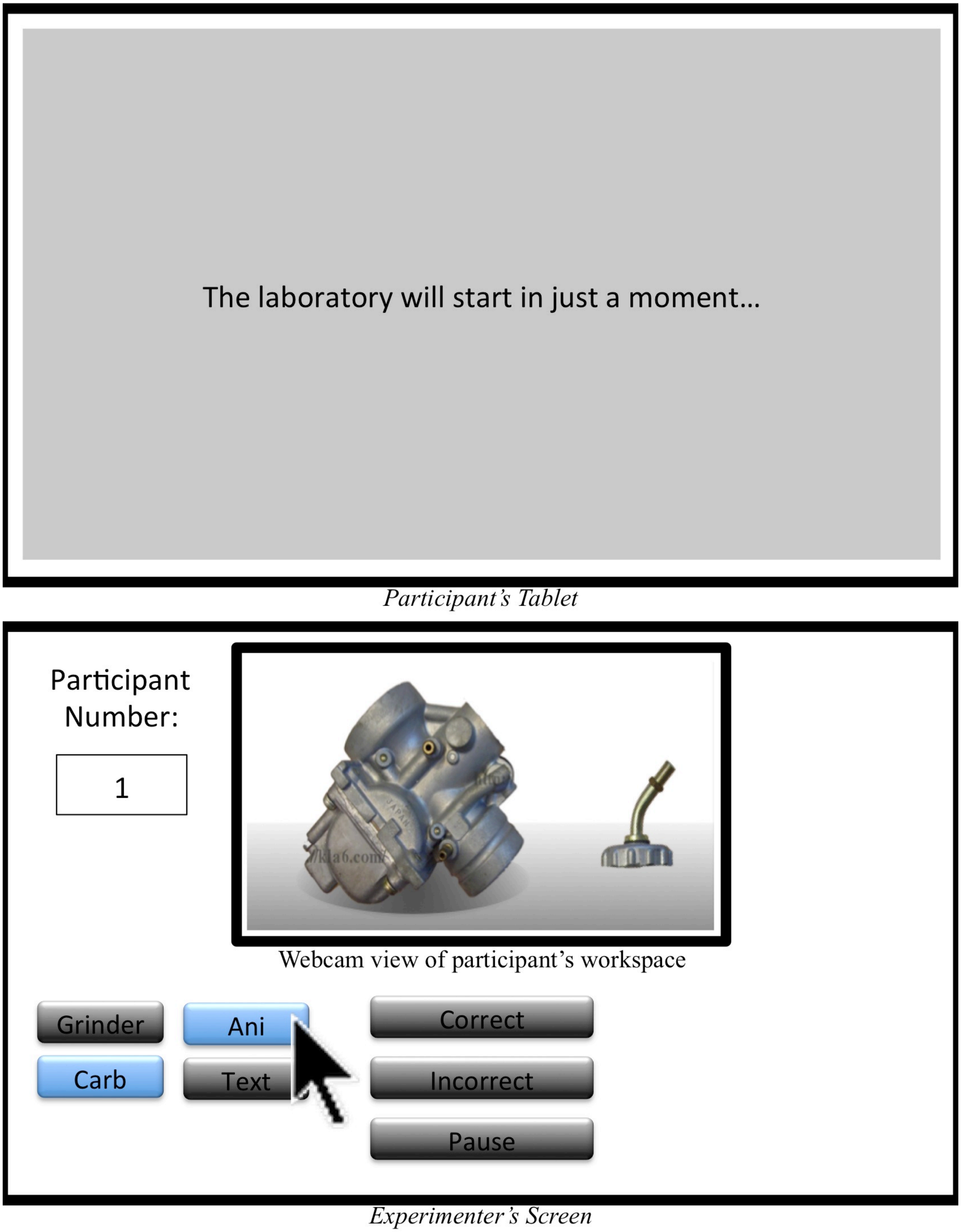


Participant's Tablet

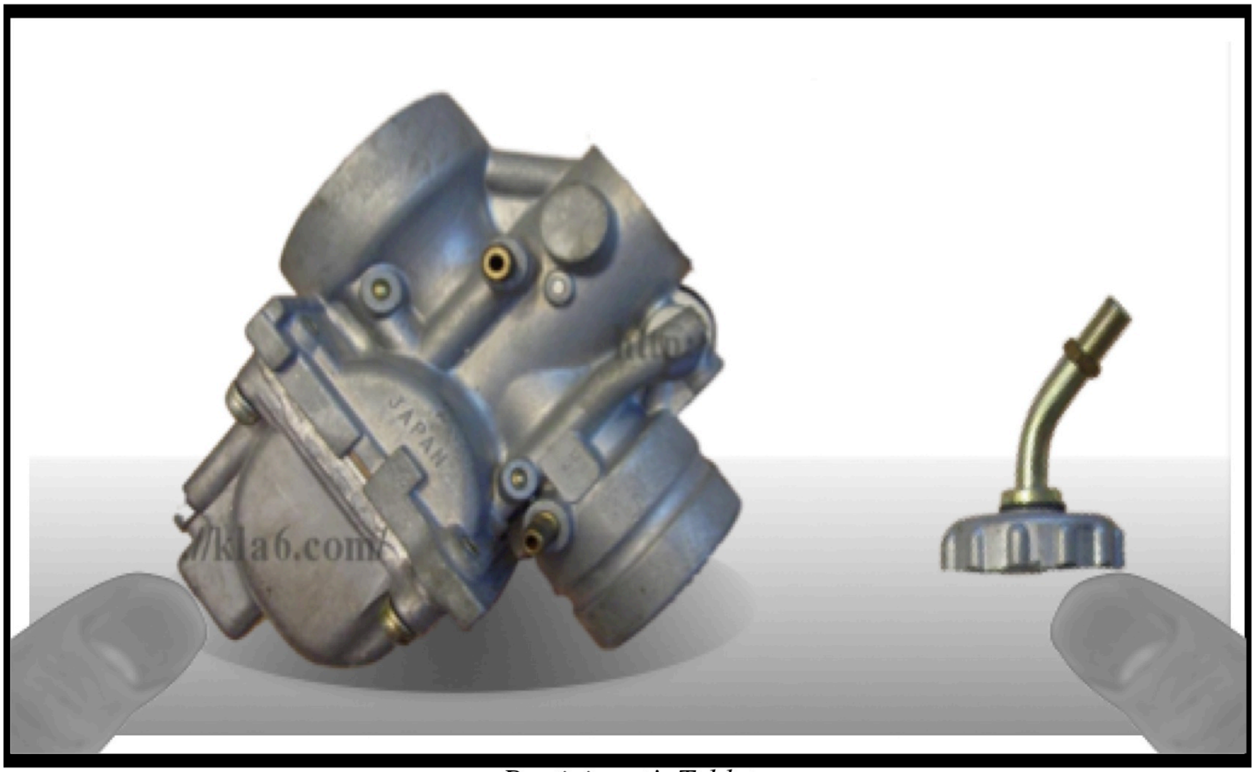


Experimenter's Screen

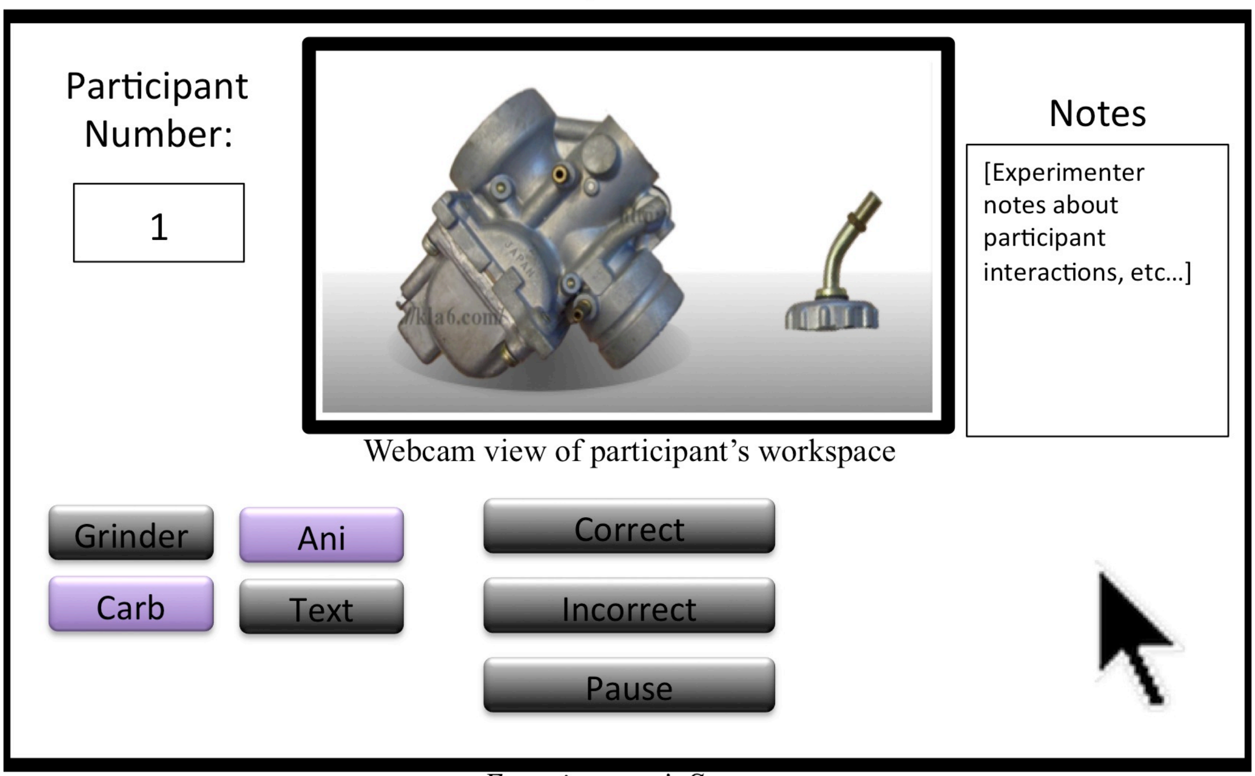
Experimenter selects appropriate instructional represetntation



Ghost thumbs prompt participant to place thumbs on screen (instruction loop starts)

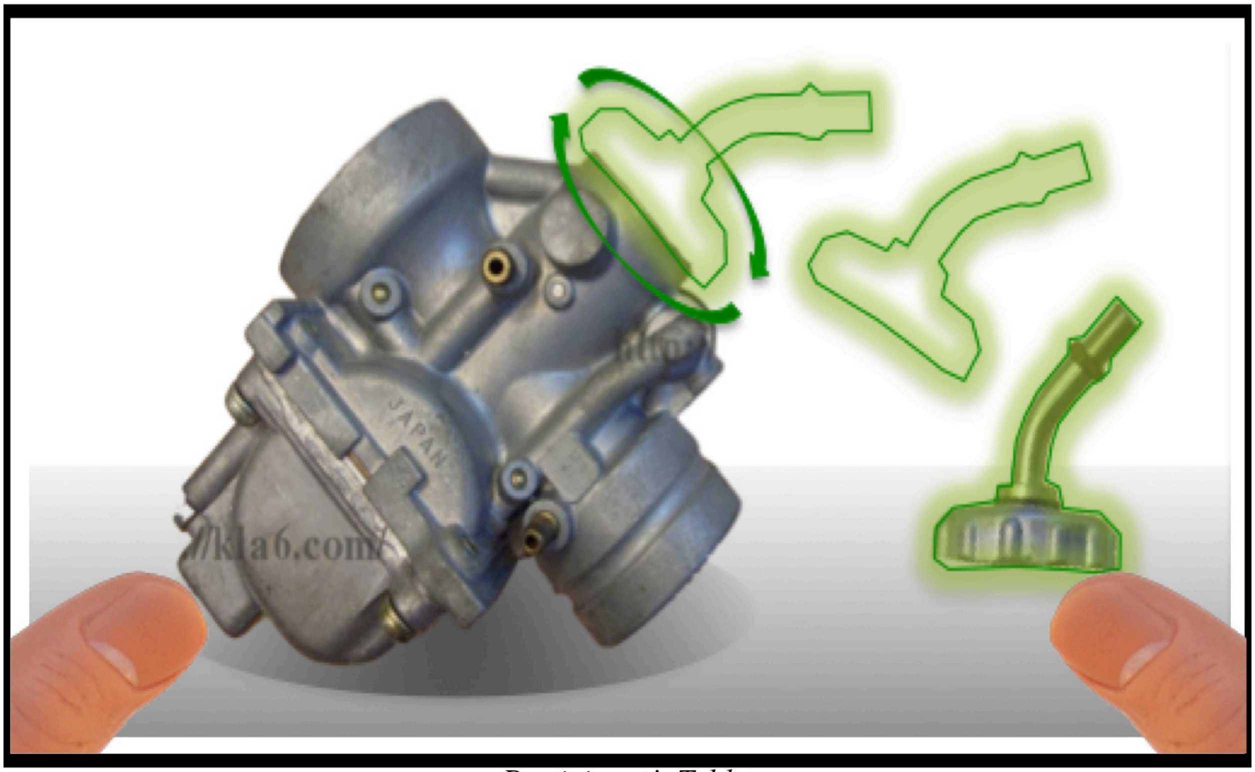


Participant's Tablet

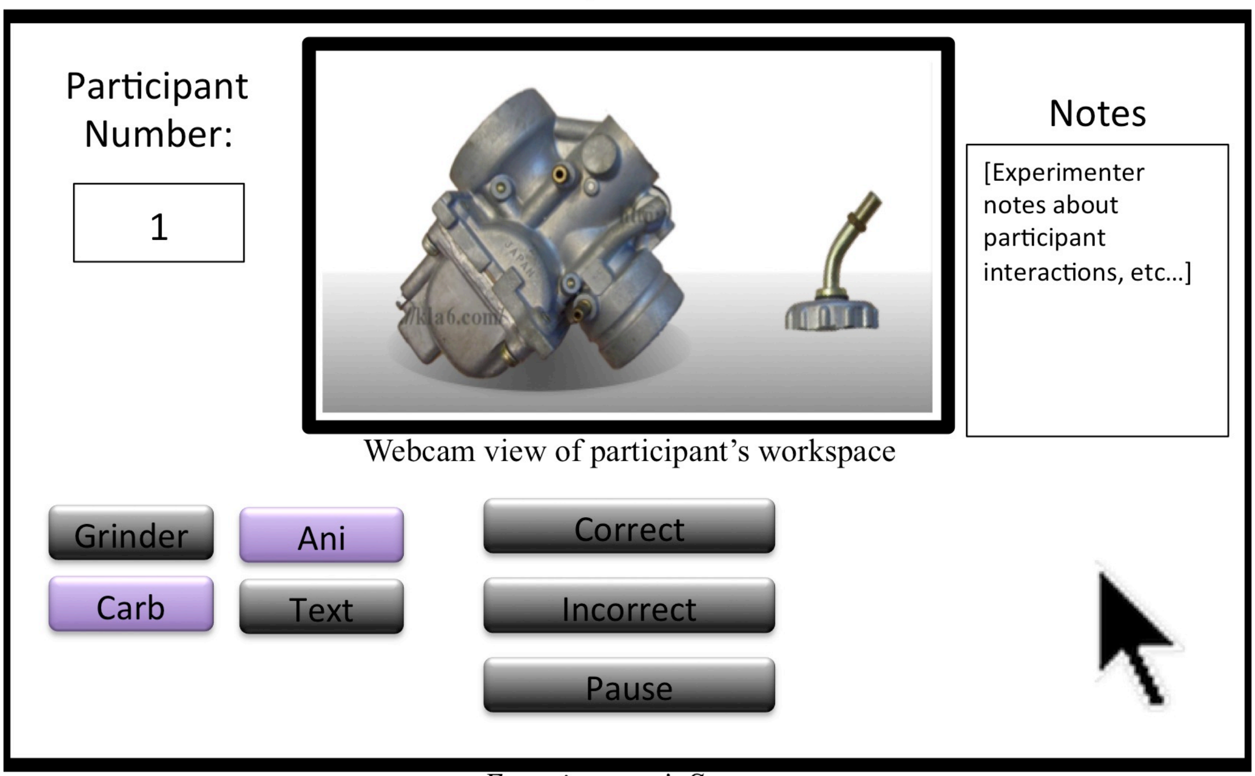


Experimenter's Screen

Instructions display only while thumbs touching screen

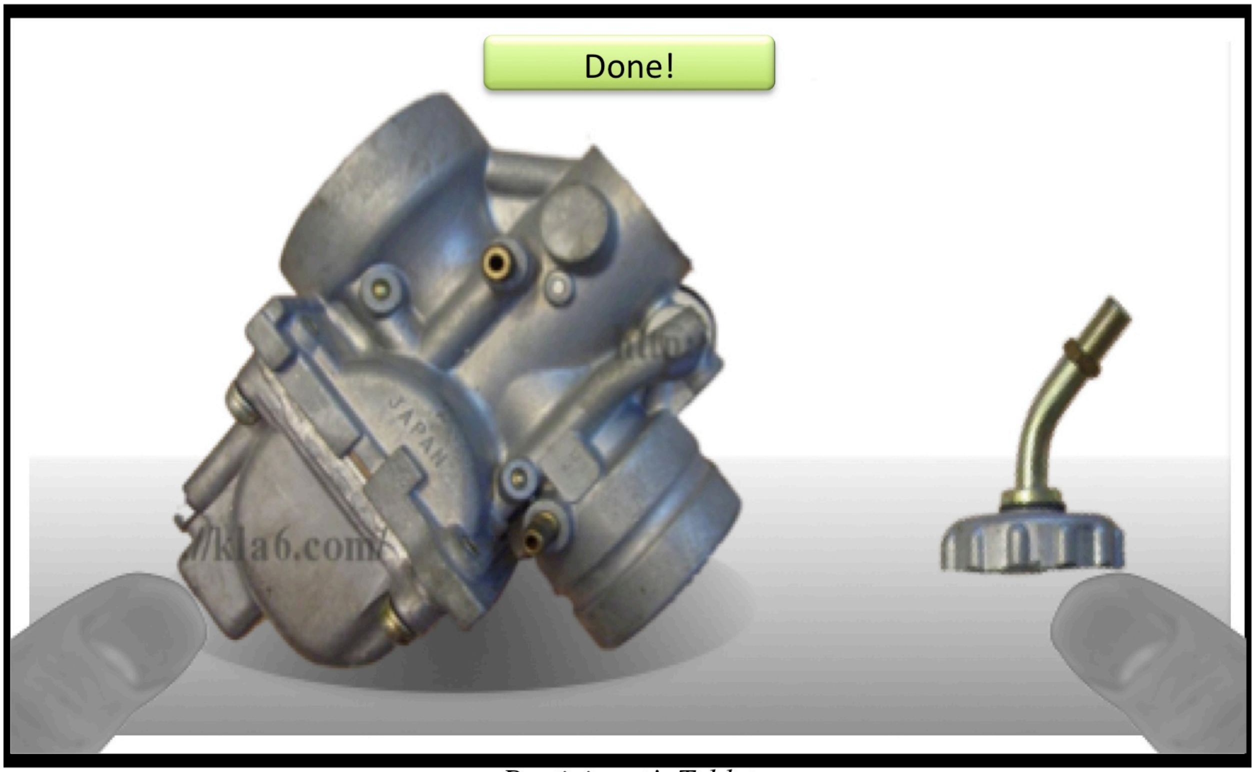


Participant's Tablet

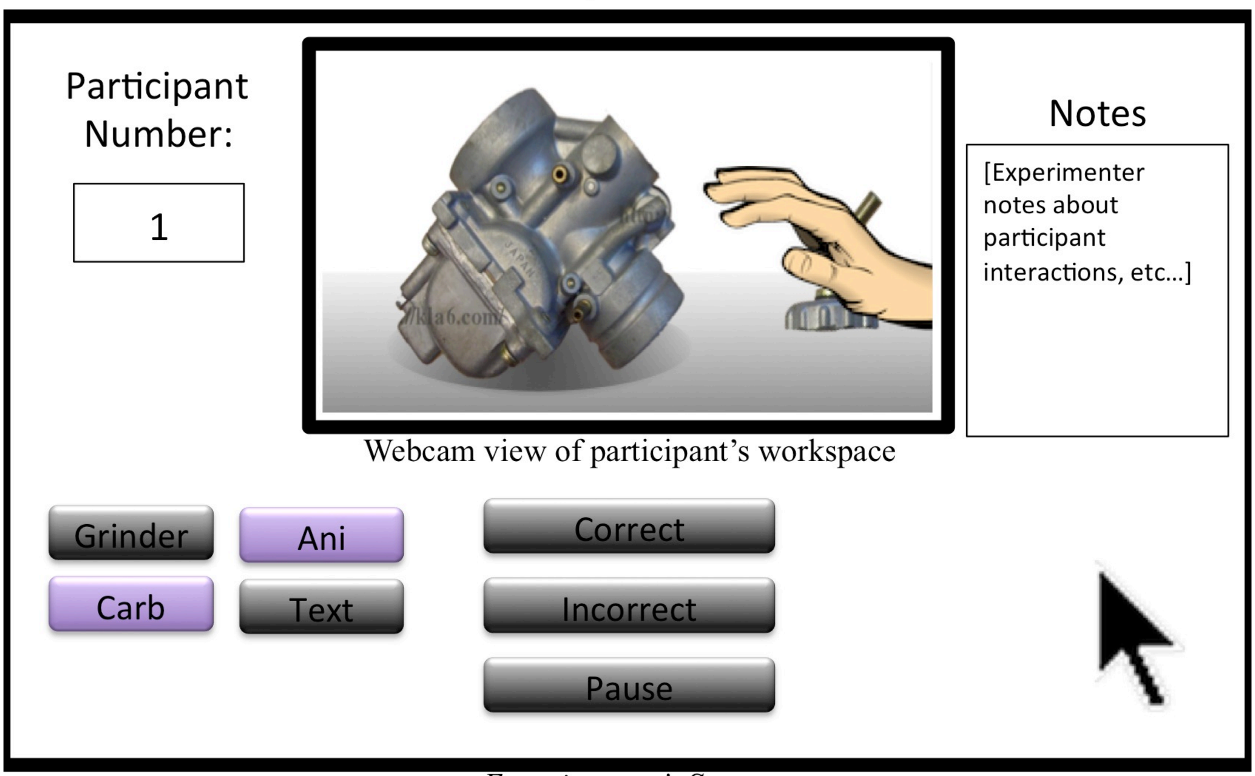


Experimenter's Screen

‘Done!’ appears, participant acts, experimenter views (instruction loop ends with ‘DONE!’)

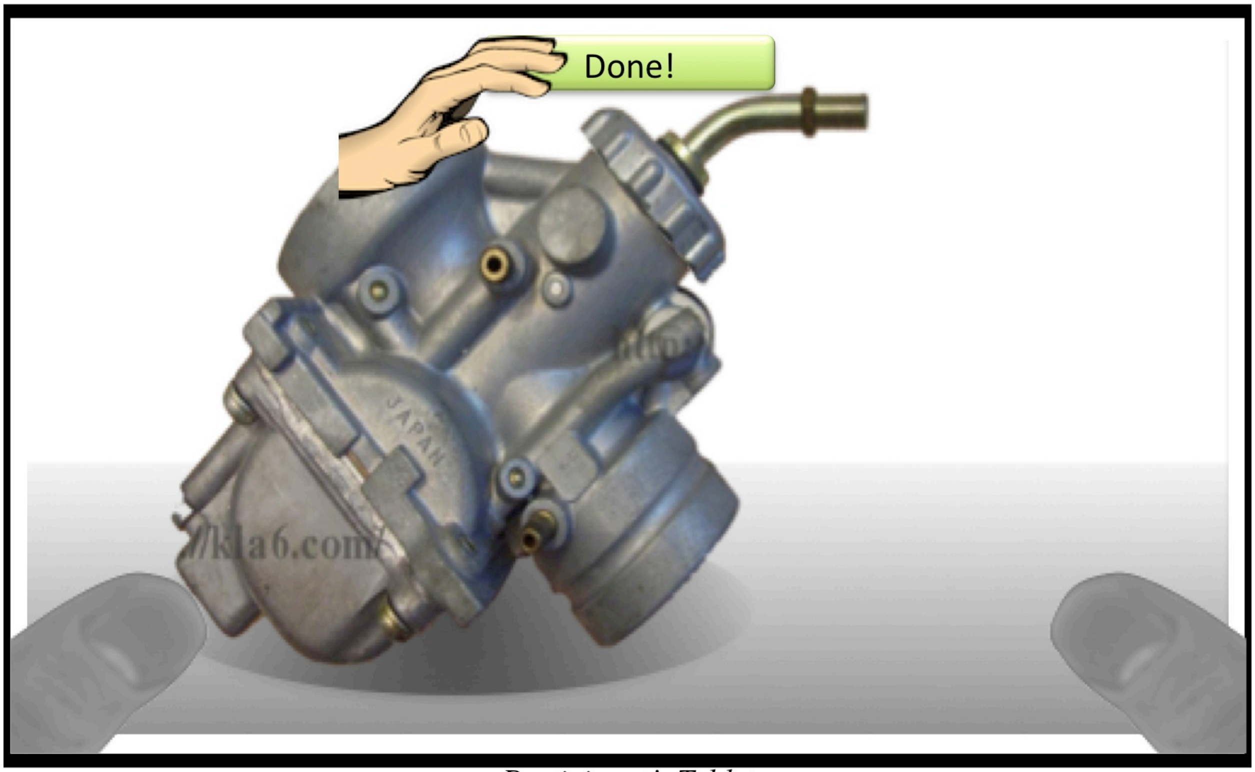


Participant’s Tablet

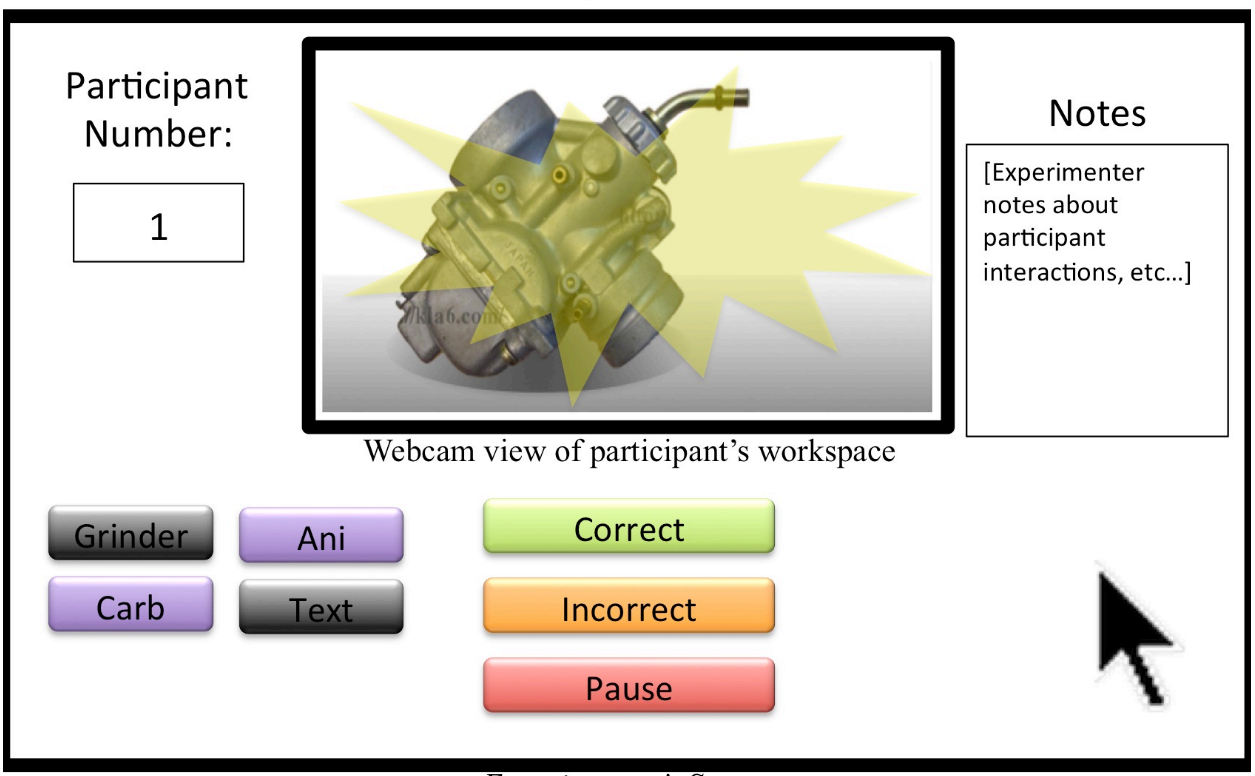


Experimenter’s Screen

Participant taps “Done!”, experimenter’s screen flashes, feedback buttons activate



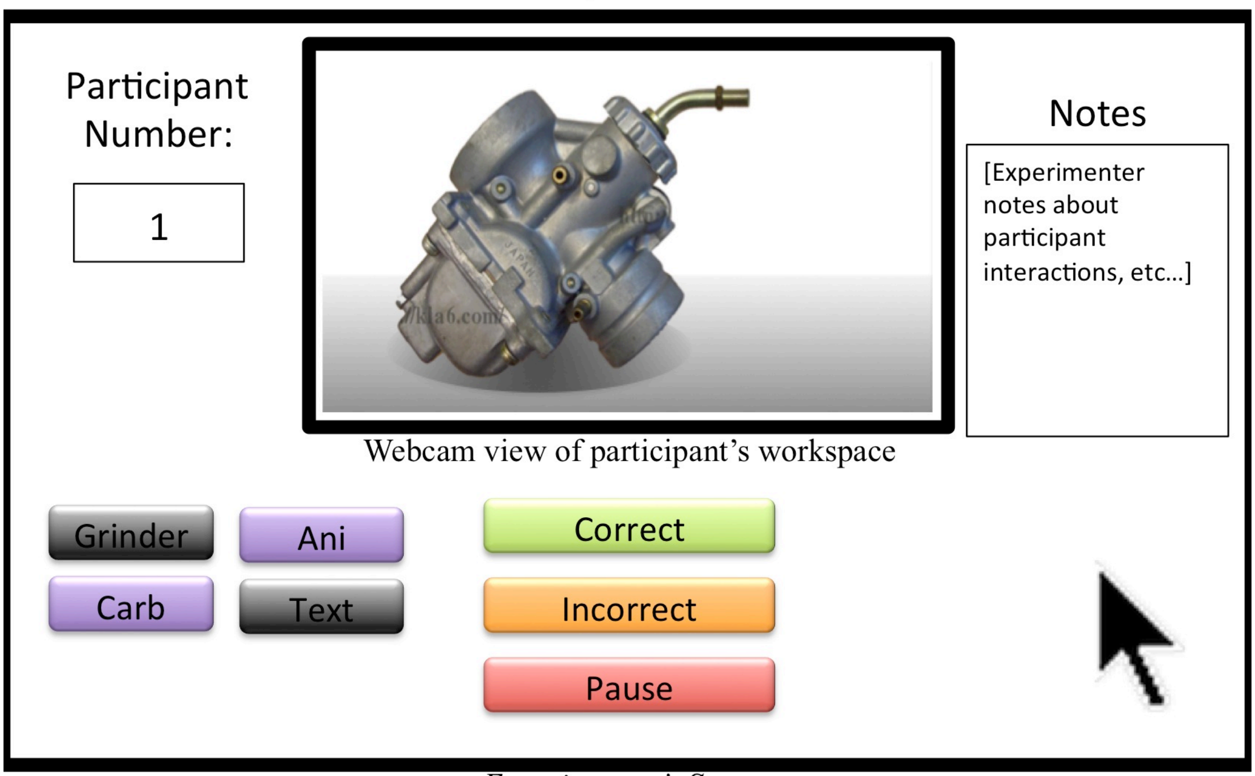
Participant’s Tablet



Experimenter’s Screen

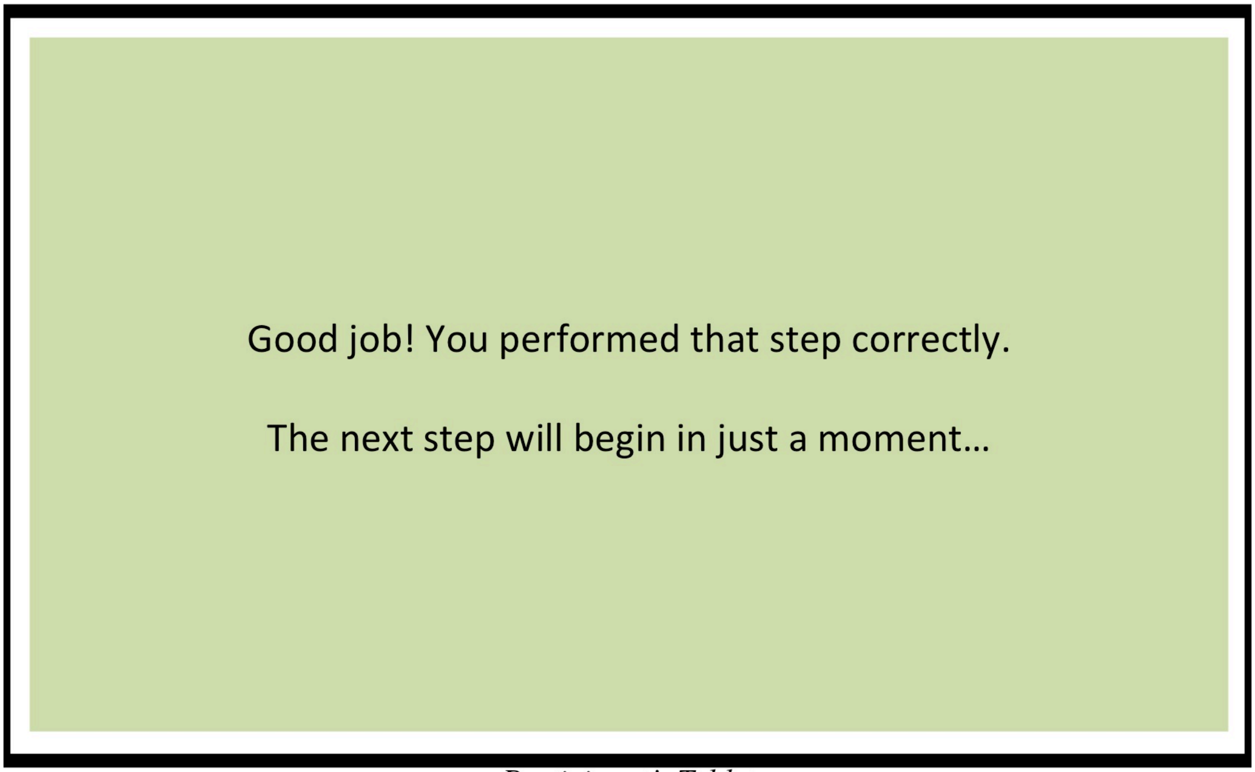
Tablet goes grey; experimenter views result, determines outcome…

Participant's Tablet

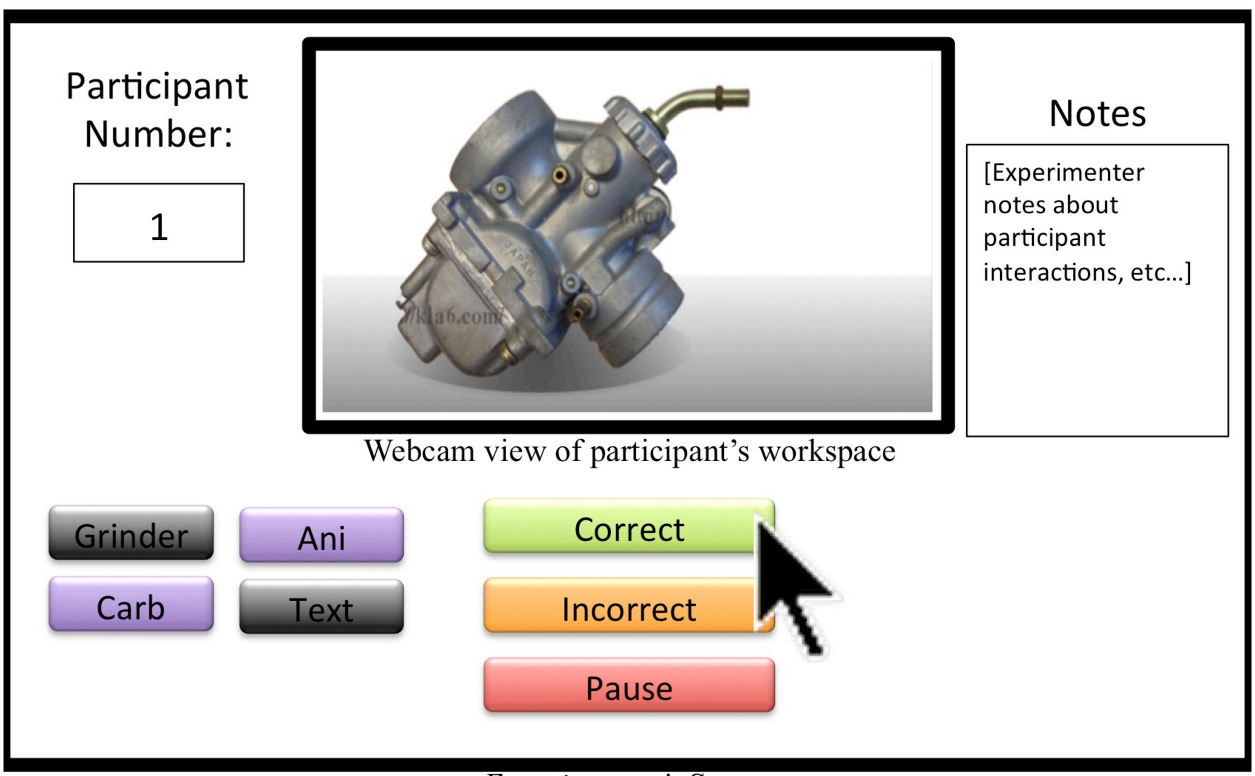


Experimenter's Screen

…participant performed step correctly, goes to next step (similar to slide 2 but for step #2)

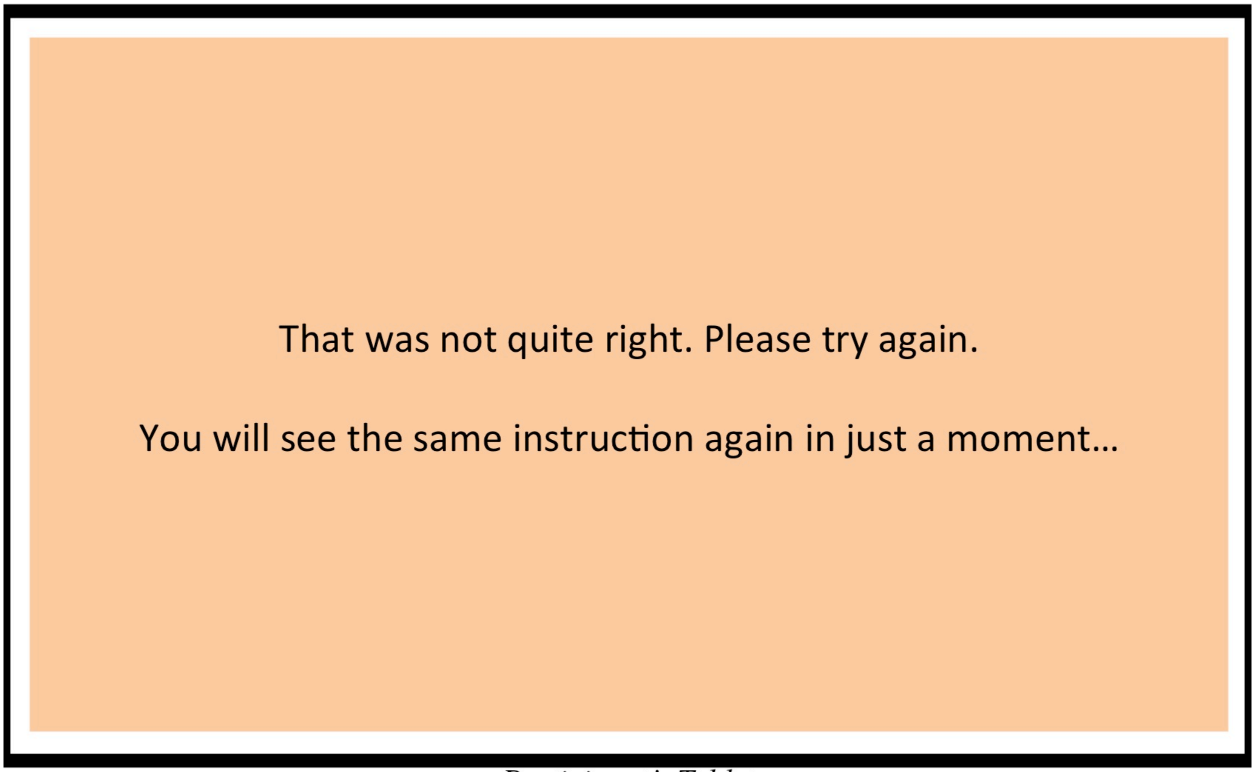


Participant's Tablet

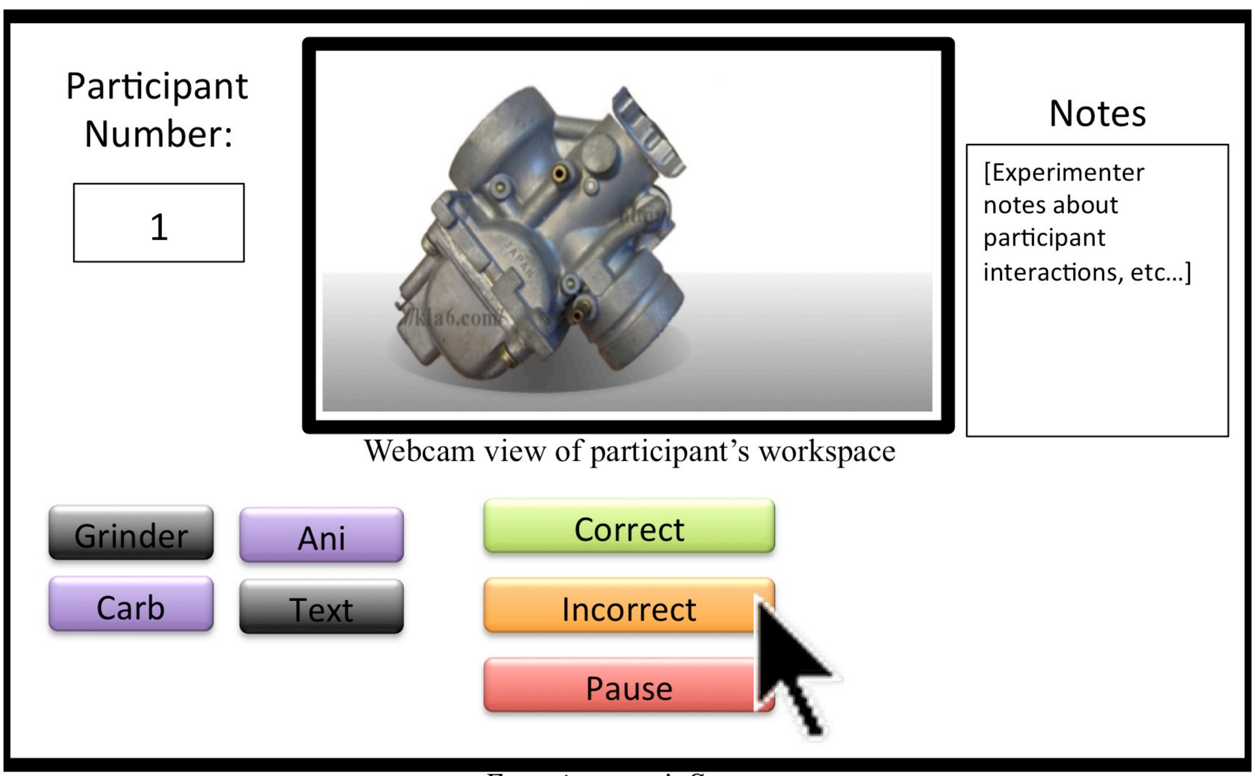


Experimenter's Screen

…participant made mistake, instruction process repeats (back to slide #2, one attempt to fix)

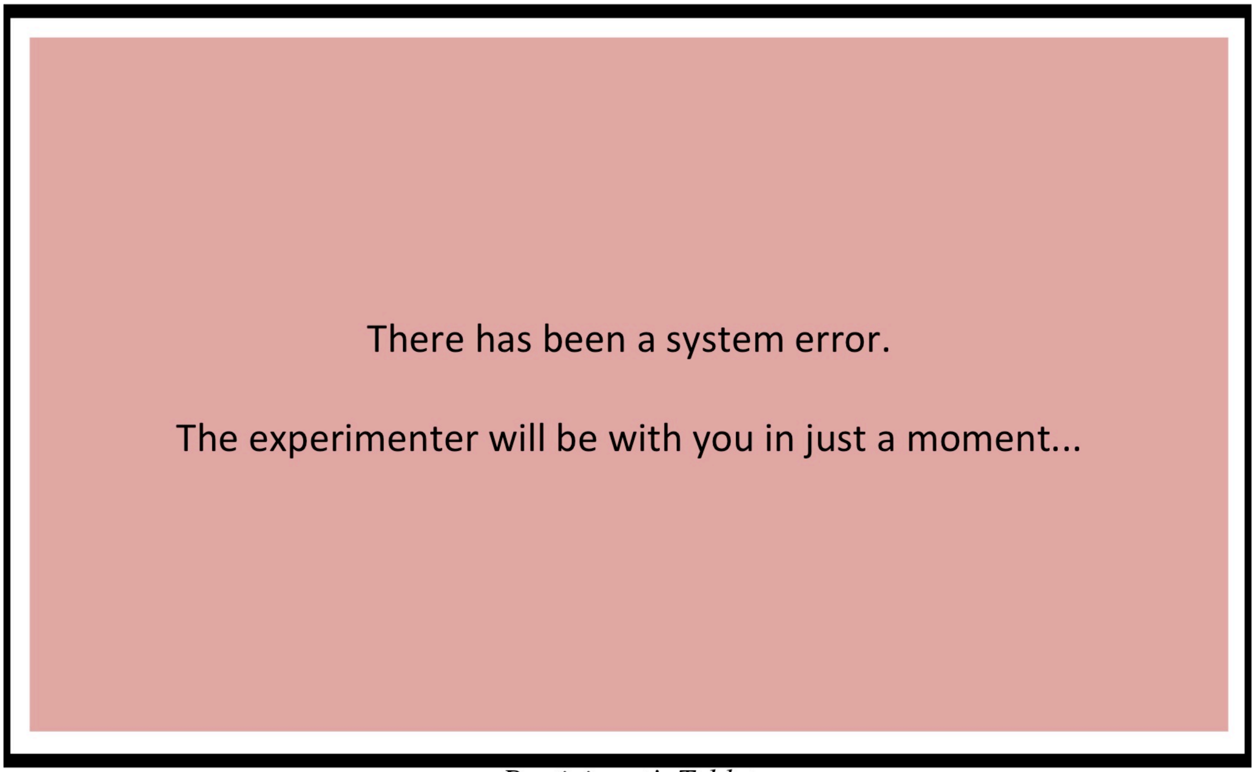


Participant's Tablet

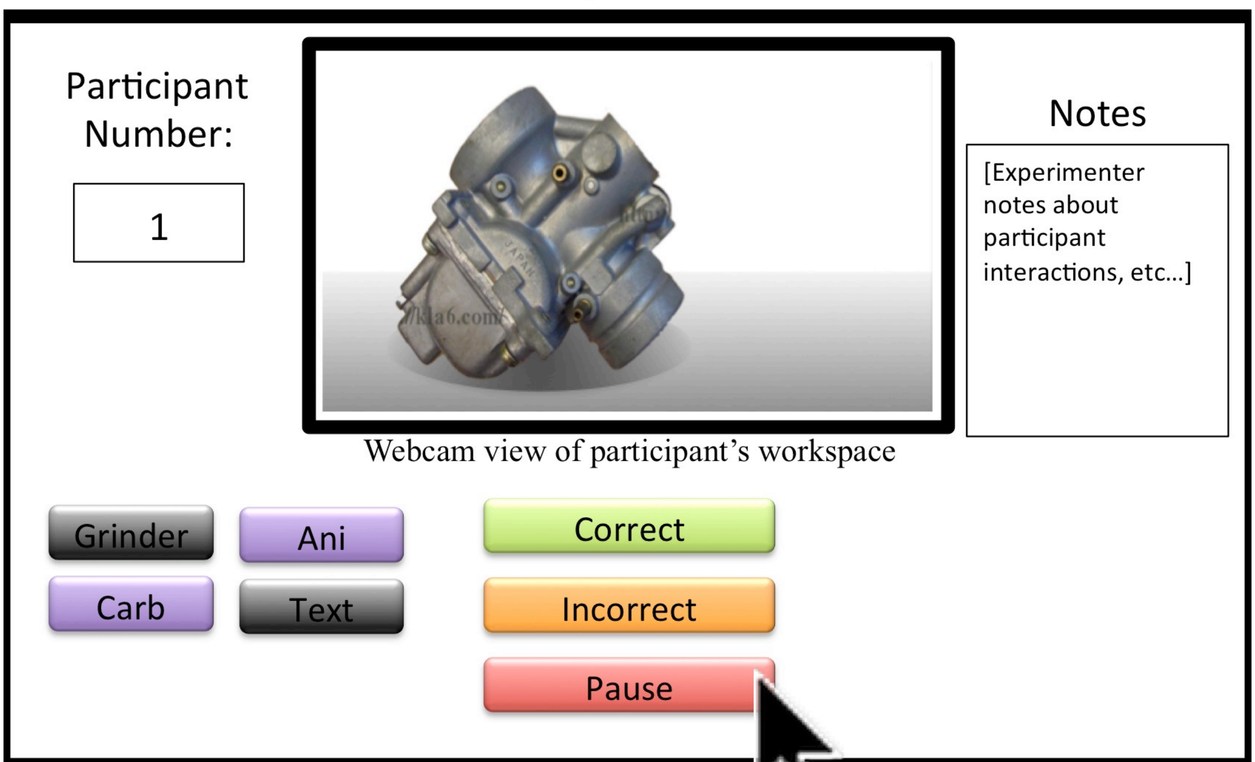


Experimenter's Screen

…experimenter fixes issue, clicks pause again (back to slide #2, one attempt to fix)



Participant's Tablet



Experimenter's Screen

REFERENCES

Adams, J. A. (1987). Historical review and appraisal of research on learning, retention, and transfer of human motor skills. *Psychological Bulletin, 101*(1), 41-74.

Ainsworth, S. (2006). Deft: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*(3), 183-198.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*(4), 612-637.

Bieger, G. R., & Glock, M. D. (1984). The information content of picture-text instructions. *Journal of Experimental Education, 53*(2), 68-76.

Bovair, S., & Kieras, D. E. (1991). Toward a model of acquiring procedures from text. In R. Barr, M. L. Kamil, P. Mosenthal & P. D. Pearson (Eds.), *Handbook of reading research, vol. 2*. White Plains, NY: Longman.

Brunyé, T. T., Taylor, H. A., & Rapp, D. N. (2008). Repetition and dual coding in procedural multimedia presentations. *Applied Cognitive Psychology, 22*(7), 877-895.

Brunyé, T. T., Taylor, H. A., Rapp, D. N., & Spiro, A. B. (2006). Learning procedures: The role of working memory in multimedia learning experiences. *Applied Cognitive Psychology, 20*(7), 917-940.

Chemero, A. (2003). An outline of a theory of affordances. *Ecological Psychology, 15*(2), 181.

Durso, F. T., & Johnson, M. K. (1979). Facilitation in naming and categorizing repeated pictures and words. *Journal of Experimental Psychology: Human Learning and Memory, 5*(5), 449-459.

Eiriksdottir, E., & Catrambone, R. (2011). Procedural instructions, principles, and examples: How to structure instructions for procedural tasks to enhance performance, learning, and transfer. *Human Factors*.

Ekstrom, R. B., French, J. W., & Harman, H. H. (1979). Cognitive factors: Their identification and replication. *Multivariate Behavioral Research Monographs*.

Ganier, F. (2001). Processing text and pictures in procedural instructions. *Information Design Journal, 10*(2), 146-153.

Ganier, F. (2004). Factors affecting the processing of procedural instructions: Implications for document design. *Transactions on Professional Communication, 47*(1), 15-26.

Ganier, F. (2012). Cognitive models of processing procedural instructions. In A. Rothkegel & S. Ruda (Eds.), *Communication on and via technology* (pp. 39-63). Berlin, Germany: de Gruyter Mouton.

Gray, W. D., & Fu, W.-T. (2001). Ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head. *Human factors in computing systems, 3*(1), 112-119.

Guthrie, J., Bennett, S., & Weber, S. (1991). Processing procedural documents: A cognitive model for following written directions. *Educational Psychology Review, 3*(3), 249-265.

Hard, B. M., Lozano, S. C., & Tversky, B. (2006). Hierarchical encoding of behavior: Translating perception into action. *Journal of Experimental Psychology: General, 135*(4), 588-608.

Hart, S. G., & Staveland, L. E. (1988). Development of nasa-tlx (task load index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload.* (pp. 139-183). Oxford England: North-Holland.

Hegarty, M., Kriz, S., & Cate, C. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition and Instruction, 21*(4), 209-249.

Heiser, J., Phan, D., Agrawala, M., Tversky, B., & Hanrahan, P. (2004). Identification and validation of cognitive design principles for automated generation of assembly instructions. *Proceedings of the working conference on Advanced visual interfaces*, 311-319.

Hmelo, C. E., & Guzdial, M. (1996). Of black and glass boxes: Scaffolding for doing and learning. *Proceedings of the international conference on Learning sciences*, 128-134.

Hochmitz, I., & Yuviler-Gavish, N. (2011). Physical fidelity versus cognitive fidelity training in procedural skills acquisition. *Human Factors*.

Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction, 17*(6), 722-738.

Jones, K. S. (2003). What is an affordance? *Ecological Psychology, 15*(2), 107.

Kintsch, W. (2008). Symbol systems and perceptual representations. In M. De Vega, A. Glenberg & A. Graesser (Eds.), *Symbols and embodiment*. Oxford: Oxford University Press.

Kissane, M., Kalyuga, S., Chandler, P., & Sweller, J. (2008). The consequences of fading instructional guidance on delayed performance: The case of financial services training. *Educational Psychology, 28*(7), 809-822.

Konz, S. A., & Dickey, G. L. (1969). Manufacturing assembly instructions: A summary. *Ergonomics, 12*(3), 369-382.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*(1), 65-100.

McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. In H. R. Brian (Ed.), *Psychology of learning and motivation* (Vol. 51, pp. 297-384): Academic Press.

Norman, D. A. (2002). *The design of everyday things*. New York, NY, USA: Basic Books.

Palmiter, S., & Elkerton, J. (1991). *An evaluation of animated demonstrations of learning computer-based tasks*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New Orleans, Louisiana, USA.

Palmiter, S., Elkerton, J., & Baggett, P. (1991). Animated demonstrations vs written instructions for learning procedural tasks: A preliminary investigation. *International Journal of Man-Machine Studies, 34*(5), 687-701.

Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *Journal of the Learning Sciences, 13*(3), 423 - 451.

Puntambekar, S., & Hübscher, R. (2005). Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed? *Educational Psychologist, 40*(1), 1-12.

Rodriguez, M. A. (2002). Development of diagrammatic procedural instructions for performing complex one-time tasks. *International Journal of Human-Computer Interaction, 14*(3-4), 405-422.

Romiszowski, A. (1999). The development of physical skills: Instruction in the psychomotor domain. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: A new paradigm of instructional theory, vol 2* (pp. 457-481). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Schnotz, W. (2002). Commentary: Towards an integrated view of learning from text and visual displays. *Educational Psychology Review, 14*(1), 101-120.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction, 13*(2), 141-156.

Turvey, M. T. (1992). Affordances and prospective control: An outline of the ontology. *Ecological Psychology, 4*(3), 173.

Watson, G., Butterfield, J., Curran, R., & Craig, C. (2010). Do dynamic work instructions provide an advantage over static instructions in a small scale assembly task? *Learning and Instruction, 20*(1), 84-93.

Wertsch, J. V. (1998). *Mind as action*. New York, NY, USA: Oxford University Press.

Yuviler-Gavish, N., Yechiam, E., & Kallai, A. (2011). Learning in multimodal training: Visual guidance can be both appealing and disadvantageous in spatial tasks. *International Journal of Human-Computer Studies, 69*(3), 113-122.

VITA

KEITH R BUJAK

BUJAK was born in Rockville, Connecticut. He attended public schools in Somers, Connecticut, received a B.S. in Mechanical Engineering and Psychology from Rensselaer Polytechnic Institute, Troy, New York in 2003 and a M.S. in Engineering Psychology from the Georgia Institute of Technology, Atlanta, GA in 2010 and continued at the Georgia Institute of Technology to pursue a doctorate in Engineering Psychology. When he is not working on his research, Mr. Bujak enjoys food and photography.